CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Research Papers in Economics

Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.sssup.it/~LEM/

LEM
Working Paper Series

Explaining the Distribution of Firms Growth Rates

Giulio BOTTAZZI*
Angelo SECCHI*

*LEM, Sant'Anna School for Avanced Studies

2005/16 **June 2005**

This paper can be considered in many respects an improved version of WP LEM 2002-14 G. Bottazzi, A. Secchi " On The Laplace Distribution of Firms Growth Rates" .

We present a new the description of the empirical results and we try to provide better justification for the theoretical assumptions constituting the base of our analysis. In the present version we introduce a new more general version of the main theorem that helps to clarify the assumption about micro-shocks distribution, the nature of the considered limits and the nature of the observed convergence.

Since the present version lacks several analysis that were performed in the aforementioned paper we decided to add the present work as a new working paper rather than a replacement of the previous one.

Explaining the Distribution of Firms Growth Rates

Giulio Bottazzi* Angelo Secchi**

Empirical analyses on aggregated datasets have revealed a common exponential behavior in the shape of the probability density of the corporate growth rates. We present clearcut evidence on this topic using disaggregated data. We explain the observed regularities proposing a model in which the firms' ability of taking up new business opportunities increases with the number of opportunities already exploited. A theoretical result is presented for the limiting case in which the number of firms and opportunities go to infinity. Moreover, using simulations, we show that even in a small industry the agreement with asymptotic results is almost complete.

1 Introduction

One of the most traditional problem in the Industrial Organization literature concerns the statistical properties of the size of firms and its dynamics.

Early investigations focused on two aspects of the general problem, namely the analysis of the size distribution and the characterization of firms growth dynamics in terms of autoregressive stochastic processes. The log-normal character of the upper tail of the size distribution was quite unanimously considered the natural benchmark. On the other hand the dynamic analysis relied on the estimate of linear models on the growth rates process in order to both verify the Gibrat hypothesis (Gibrat, 1931) of random-walk growth and to find possible violations (in the enormous body of contributions see for instance Dunne et al. (1988); Evans (1987a); Hall (1987)).

These early works were conducted over datasets at a high level of aggregation, typically including large firms operating in very different sectors. For instance, Hart and Prais (1956) studied the dynamic of the whole U.K. manufacturing industry, while Simon and Bonini (1958) and Hall (1987) explored the size and

*Scuola Superiore S.Anna; bottazzi@sssup.it.
**Scuola Superiore S.Anna; secchi@sssup.it.
Support by the Italian Ministry of Education, University and Research MIUR (grant n. A.AMCE.E4002GD) and the S.Anna School of Advanced Studies, Pisa (grant n. E6003GB) is gratefully acknowledged. The authors thank M. Anoufriev, C. Castaldi, G. Dosi, S. Modica and M. Sylos Labini for helpful discussions and A. Pakes and two anonymous referees for insightful comments. The research would not have been possible without the precious help of the Italian Statistical Office (ISTAT) and in particular of Roberto Monducci. The usual disclaimers apply

growth process of the manufacturing firms of the U.S. economy, across all the sectors. A common source of problems in considering such aggregate data is the possibility of introducing statistical regularities that are simply the result of the aggregation process and, at the same time, concealing the true properties of the dynamics of business firms that are active in specific sectors. Indeed Hymer and Pashigian (1962), analyzing more disaggregated data, find a high heterogeneity in firms size distributions across different sectors. They conclude that it is quite unclear whether any "stylized fact" regarding the size distribution actually exists. As far as the validity of the Gibrat's hypothesis is concerned, the conclusions of these works are variegated, if not contradictory (see Singh and Whittington (1975) for an early sectoral analysis and the critical reviews in Sutton (1997) and Lotti et al. (2003)).

Moving from the foregoing traditional econometric issues, a new strand of analysis recently emerged proposing a more complex statistical characterization of firms growth dynamics. Following these lines of research this paper, extending preliminary results reported in Bottazzi and Secchi (2003), analyses the growth rates distribution of business firms in the Italian manufacturing industry using data disaggregated by sectors. The results are clearcutting: the growth rates probability density, in all the sectors under study, possesses the same symmetric exponential character that, when plotted on log scale, emerges as a sort of tent-like shape. The same tent-shape characterizes growth rates density in U.S. manufacturing industry (Stanley et al., 1996) and in the world-wide pharmaceutical industry (Bottazzi et al., 2001).

The robustness of this empirical finding constitutes an interesting theoretical issue unexplained by the few standard models present in the literature. In our opinion, the reason for that can be traced back to the presence, in those models, of noticeable weaknesses. First of all, from the seminal work of Gibrat (Gibrat, 1931) to the more recent contributions of Geroski (2000) and Amaral et al. (2001), many models do not assume any interdependence between the histories of different firms. The dynamics of each firm is a stochastic process, encompassing growth, diversification, entry and exit, that, nevertheless, does not take into consideration the behavior of the other firms. Each firm acts as if it was a monopolist in a sector whose dynamics can be represented simply with an exogenous expansion (or contraction) of demand. A different kind of models, originally proposed by Ijiri and Simon (1977) and later reconsidered by Sutton (1998) make the assumption that there is a finite set of pre-existing growth opportunities (or equivalently, a constant arrival of new opportunities) and that firms growth process is conditioned by the number of opportunities they are able to take up. Roughly speaking, one could say that these models, generically known as "islands models", try to provide a first account of the competitive behavior based on the idea that firms need to seize the most out of a scarce resource. Nevertheless, these models fail to explain the empirical shape of the growth rates density.

A complementary stream of literature, encompassing a vast body of models, has proposed the inclusion of a competitive dimension in the description of industrial dynamics based on diverse rationality and informational assumptions concerning the behavior of economic agents. Consider, for instance, the model based on the notion of "Schumpeterian competition" in Nelson and Winter (1978), the Bayesian learning model in Jovanovic (1982) or the model of research and exploration in Ericson and Pakes (1995). Even if these models are both successful in bringing a more plausible microeconomic foundation in the description of business firms dynamics, and helpful in deriving clear empirical implications on the dynamics of single firms and on the structure of the whole industry (c.f. for instance Evans (1987b) and Pakes and Ericson (1995)), they do not focus on the specific issue of the shape of the growth rates distribution.

In the present paper we build a simple mechanism of firm dynamics where a stylized idea of competition is introduced. Nonetheless, rephrasing Nelson and Winter (1978), in our model luck is the principal factor that finally distinguishes winners from losers among the contenders. Even if we are well aware of the need of a more structural approach in the development of detailed models aimed at the description of particular industries, we take here a different perspective inspired by Simon's tradition aiming at both simplicity and generality. We introduce a stochastic description where each firm is considered a different realization of the same process. This process represents a simple generalization of already existing island models. Similarly to what happens in these models, the symmetry is broken at the aggregate level: the total growth of the whole population of firms is bounded by a finite set of sector-specific opportunities.

The novelty resides in the way in which we describe the random distribution of opportunities among firms. In the existing formulations (Ijiri and Simon, 1977; Sutton, 1998) the assignment procedure with which market opportunities are distributed carries no history and at each competitive round each firm possesses the same probability of seizing them. Conversely, our assignment procedure allows to represent self-reinforcing mechanisms whereby the probability for a given firm to take up a new opportunity positively depends on the number of opportunities already taken up.

The remainder of this paper is organized as follows. After a brief description of our data, in Section 3 we report the results of our empirical investigation. Section 4 proposes a new stochastic model of firm growth while Section 5 continues its analysis and compares its features with the empirical findings. In this Section we also present a formal result that ensures generality and robustness to our model. Section 6 draws some conclusions and briefly comments on the need for further theoretical research.

2 Data Description

This research draws upon the MICRO.1 databank developed by the Italian Statistical Office (ISTAT)[1]. MICRO.1 contains balance sheets entries of a panel of several thousands of Italian firms, over around a decade. In MICRO.1 only firms with 20 or more employees are considered and different businesses inside the same firm are assigned to the firm primary activity[2]

Since the panel is open, due to entry, exit, fluctuations around the 20 employees threshold and variability in response rates, we consider a balanced panel composed only by the firms that are present both at the beginning and at the end of our window of observation[3]. For statistical reliability we restrict our analysis to the period $1989 - 1996$ and to the sectors with more than 44 firms, reducing the number of sectors under study from 97 to 55. The choice to limit the analysis to a balanced panel containing only the largest sectors[4] reduces the total number of firms under study from around 36000 to around 8400.

In this work we are exclusively interested in the process of *internal* growth, as opposed to the growth due to mergers, acquisitions and divestments. In order to control for these phenomena we build "super-firms" which account throughout the period for the union of the entities which undertake such changes. So, for example, if two firms merged at some time, we consider them merged throughout the whole period. Conversely, if a firm is spun off from another one, we "re-merge" them starting from the separation period [5]. This "re-merging" procedure affects less then the 15% of the whole population of firms. After the application of this procedure, we end up with a sample of 8091 super-firms observed for 8 years.

3 Empirical evidence

Some years ago in a series of papers based on the COMPUSTAT database Stanley et al. (1996) and Amaral et al. (1997) analyzed the probability distribution of the (log) growth rates of publicly traded U.S. manufacturing firms. These studies were performed using observations in the time frame 1974-93 and on companies with primary activity belonging to the SIC code range 2000-3999. Different lines of business inside the same multi-activity firm were completely aggregated. According to these analyses, the firm growth rate g, when one considers the aggregate distribution across all the sectors, appears to robustly display, on a log scale, a

[1]The database has been made available to our team under the mandatory condition of censorship of any individual information.

[2]This operation is performed directly by ISTAT; hence we do not have specialization ratios.

[3]We are aware that this procedure could introduce a selection bias due to fluctuation around the threshold of 20 employees. However consider that we use Sales to proxy firm size while the inferior threshold is defined in terms of number of employees. This should reduce the severity of the bias.

[4]We replicated our analysis also on the unbalanced database, obtaining very similar results. For brevity, we do not report here this analysis. Results are available upon request.

[5]For more details on this database and on the variables used in this paper see Bottazzi et al. (2002).

characteristic tent-shape probability density. Hence a Laplace (symmetric exponential) functional form

$$f_{\mathrm{L}}(g;\mu,a) = \frac{1}{2a} e^{-\frac{|g-\mu|}{a}} \tag{1}$$

was proposed in order to describe the empirical observations. More recently Bottazzi et al. (2001) found the same characteristic shape for the empirical density of the growth rates of the largest worldwide companies in the pharmaceutical industry.

The similarity across these early studies naturally leads to the question of how general this tent-shape character is when different industries or countries are taken in consideration. Moreover, these studies where focused on very large multi-plants and/or multinational companies and, in particular for the COMPUSTAT based analysis, data were aggregated across many distinct sectors. Hence a further possible issue concerns the robustness of this finding when smaller firms and disaggregated data are analyzed. In the present section we address these two issues. Our study of the MICRO.1 dataset, that includes a large part of the Italian manufacturing industry, adds new evidence to the original finding. The analysis is conducted sector by sector in order to check to what extent the mentioned finding survives at a more disaggregated level.

In what follows we use total sales as a definition of firms' size. Let $S_{i,j}(t)$ represents the sales of the i-th firm, belonging to the j-th sector, at time t. Here $j \in \{1,\ldots,55\}$ and if N_j is the number of firms in the j-th sector, one has $i \in \{1,\ldots,N_j\}$. In order to eliminate possible trends, both sector specific and industry-wide, we consider the normalized (log) sales

$$s_{i,j}(t) = \log(S_{i,j}(t)) - \frac{1}{N_j}\sum_{i=1}^{N_j}\log(S_{i,j}(t)) \tag{2}$$

subtracting from the (log) size of each firm the average (log) size of all the firms operating in the same sector. The (log) growth rate is then defined according to:

$$g_{i,j}(t) = s_{i,j}(t+1) - s_{i,j}(t) \quad . \tag{3}$$

Notice that from (2) the distribution of the g's is by construction centered around 0 for any t.

As a first qualitative investigation one can simply plot the observed densities for different sectors. Fig. 1 and Fig. 2 show the growth rates densities for six different three digit sectors chosen because both numerous and structurally diverse. The activity indeed ranges from footwear production to the treatment of metals for industrial purposes. All the 7 years of data are pooled together under the assumption of stationarity of the growth process [6]. For each sector the Laplace density estimated via maximum likelihood is also shown. As can be seen, these fitted densities describe the observations well.

In order to quantify the agreement with the Laplace and to give a synthetic account of its robustness and generality in describing empirical densities we follow a parametric approach. We consider a flexible family of probability densities, known as the Subbotin family (Subbotin, 1923), that includes as a particular case the Laplace. The Subbotin density, centered in $g = 0$, is characterized by 2 parameters: a scale parameter a and a shape parameter b. Its functional form reads:

$$f_{\mathrm{S}}(g) = \frac{1}{2ab^{1/b}\Gamma(1/b+1)}\, e^{-\frac{1}{b}\left|\frac{g}{a}\right|^{b}} \tag{4}$$

where $\Gamma(x)$ is the Gamma function. The lower is the shape parameter b, the fatter are the density tails. For $b < 2$ the density is leptokurtic and is platikurtic for $b > 2$. It is immediate to check that for $b = 2$ this density reduces to a Gaussian and for $b = 1$ to a Laplace (symmetric exponential). For each sector we compute the density that best fits the data among those belonging to this family. We estimate the a and b parameters for each sector maximizing the likelihood of observations.

The binned empirical density of the b parameter estimates over the 55 sectors is reported in Fig. 3. The values for specific sectors can be read from Table 1 together with the Cramer-Rao standard errors obtained from the inverse information matrix (Agrò, 1995). Considering the 95.7% significance level defined by the two standard errors threshold, only 15 sectors on 55 possess values of b that are significantly different from 1. Even when this difference result significant, its absolute size is small: only 4 sectors out of 55 possess values of b that are significantly outside the interval $[.9, 1.1]$.

In Fig. 5 we report the "aggregate" empirical growth rates density, obtained pooling together the observation from the 55 sector under study. In the case of 1-year lag growth rates, a clear symmetric exponential shape appears. The maximum likelihood estimation of the Subbotin distribution on the aggregated data provides a value $b = 0.965$ with a standard error of 0.007, very similar to the Laplace value $b = 1$. This result is in perfect agreement to what found by Stanley et al. (1996) on the COMPUSTAT database and we can conclude that the tent-shape characterizes the growth rates density both at aggregate and disaggregate level. Since the mixture of Laplace densities with heterogeneous variances would approximate a Gaussian distribution, one could expect that the peculiar Laplace shape, when present at the level of single sectors, would tend to disappear in the aggregate. The apparent lack of this effect in our data is due to the fact that the sectoral growth rate

[6] In performing this pooling we are assuming that the conditional distribution of firm growth rates is independent from the size of the firm. Even if this is not generally the case we have checked that this requirement is fulfilled across all sectors of our database. For a discussion see Bottazzi et al. (2003).

densities are similar not only in terms of their tail behavior, described by the parameter b, but also in terms of their typical “width”, captured by the parameter a. Indeed, also the empirical density of the ”scale” parameters a, estimated via likelihood maximization and reported both Fig. 4 and in Table 1 (together with standard errors), possesses a remarkably narrow support. This evidence suggests a quite strong similarity among the growth rates densities in different sectors and, consequently, it helps to preserve the Laplace shape even when different sectors are considered together.

We conclude our analysis of the firms growth rates by looking at their structure on a longer, multi-year, time horizon. In line with previous notation we define the growth rate on a T year period as:

$$g_{i,j}(t;T) = s_{i,j}(t+T) - s_{i,j}(t) \quad . \tag{5}$$

When $T = 1$, (5) reduces to the one year growth rates defined in (3). Using again maximum likelihood estimation one can compute the value of the a and b parameters in each sector at different T. As can be seen from Fig. 6 the average value of the b parameters across all the sectors, that is near to 1 when $T = 1$, steadily increases for longer intervals. This implies that the typical shape of the growth rates density becomes more similar to a Gaussian when longer time horizons are considered. An example of this effect is shown in Fig. 5 where the aggregate growth rates density is reported in the case of a time lag of 7 years ($T = 7$) together with the best Subbotin fit which provides a value of $b = 1.243$ with a standard error of 0.027 lying between the Laplace $b = 1$ and the Gaussian $b = 2$ case. This phenomenon might be considered natural if the firm growth shocks relative to different years were independent and, consequently, the progressive normalization of the growth rate density were an effect of the Central Limit Theorem (CLT). Notice, however, that the slope of the curve in Fig. 6 seems to decrease rapidly as T increases, suggesting that the asymptotic value of b can be quite below the expected value of 2. The time horizon of our database is however too short to allow a reliable discussion of this point. As one consider longer time lag, the number of available observations decreases and the statistics become so noisy that it is impossible to conclude if some further effect, apart the CLT, is at work here. Moreover, the heterogeneity of the autocorrelation coefficients of the growth rates in different sectors (see Table 1) tends to complicate the matter.

The empirical findings of this section can be summarized as follows. First, the Laplace density constitutes a good approximation of the observed densities not only when large firms and/or aggregated data are considered, but also for medium sized firms and at a disaggregated level. Second, this characteristic shape tends to disappear when longer time horizons are considered, as suggested by the Central Limit Theorem under the hypothesis of independent growth events.

Our investigation is then arrived at the end of its first stage: we found a simple generalization (the Laplace distribution) that is able to describe an empirical fact (the tent-shape of the empirical growth rates density) with, in our opinion, a good degree of approximation. In the very spirit of the Simonian tradition (Simon, 1968), our second step will be to propose a possible explanation for this stylized fact and identify the conditions under which the deviations of the empirical observations from the proposed explanation my be expected to decrease.

4 A model of firm growth

In the literature about the stochastic models of firms growth dynamics, there exists a well established tradition describing the modification of firm size, over a given period of time, as the cumulative effect of a number of different "size" shocks generated by the diverse accidents that affected the firm history in that period (see, among many contributors, Kalecki (1945); Ijiri and Simon (1977); Steindl (1965) and more recently Amaral et al. (2001); Geroski (2000); Sutton (1998)). Since these models are usually described in terms of multiplicative processes, it is natural to use logarithmic quantities to describe the "status" of a given firm. Consider a firm i and let $s_i(t)$ be its (log) size at time t. One can write

$$g_i(t;T) = s_i(t+T) - s_i(t) = \sum_{j=1}^{G_i(t;T)} x_j(t) \tag{6}$$

where the firm growth in the period $[t, t+T]$ is described as a sum of $G_i(t;T)$ "shocks" each one having a finite effect x on firm size. In empirical studies, the time lag T can range from 3 months for quarterly data, to $30-50$ years for the longest databases. In the oldest model of Gibrat (Gibrat, 1931) the shocks x's are assumed independent realizations of the same random variables $\mathbf{x}$, so that the firms growth is described as a geometric Brownian motion. The growth rates associated to different non-overlapping time periods are independent and when the number of shocks $G_i(t;T)$ increases, the rate of growth $\mathbf{g}_i(t;T)$ tends, for the Central Limit Theorem, toward a normal distribution.

As we showed in the previous Section this is not the case in the real world: in three very different databases, at least when yearly data are considered, a Laplace distribution fits the data much better than a Gaussian.

Since the Gibrat's model cannot yield an equilibrium distribution of the growth rates that resembles the observed one, we are led to conclude that some of the assumptions adopted are not appropriate.

Probably the most noticeable drawback of the Gibrat's idea resides in the implicit assumption that companies growth process are completely independent. This is equivalent to assume the absence of any form of competition, even among firms operating in the same sector and selling on the same market. To this respect, a different theoretical tradition, dating back to the early work of Simon and recently renewed by Sutton, aims to introduce in the family of Gibrat-type stochastic models of growth a stylized description of competition and entry dynamics.

These "islands models" postulate the existence of a finite number of business opportunities available to firms. All the firms, operating in a number of independent sub-markets (islands), take up the available opportunities and their growth process is measured by the number of opportunities they end up with. These opportunities represent all sorts of "accidents" that can plausibly affect the history of a business firm: the exploitation of technological novelties, the reaction to demand shocks and the effects of managerial reorganizations. The departure of these models from the Gibrat tradition is twofold. First, even if each business opportunity concerns only one firm, the symmetry of the growth process of different firms is broken, in the aggregate, by the fact that the business opportunities are limited. Second, there is always a finite probability that business opportunities are taken up by new firms. In these models the constant G_i is reinterpreted as a stochastic variable $\mathbf{G}_i$, representing the outcome of a random assignment procedure of business opportunities among incumbent and entrant firms.

It turns out that even the "island models" (both the original version by Simon and the most recent refinement by Sutton) fail to account for the observed tent-shaped density of growth rates. Indeed, if one switches off the entry dynamics, as we did in the empirical investigations presented in the previous Section, these models again generate a Gaussian growth density. This stems from the assumed equiprobability of incumbent firms to capture new business opportunities when the process is described in terms of logarithm. In this case the unconditional distribution of $\mathbf{G}_i$ for a given firm is binomial; so that, in the limit of many small opportunities, one obtains again, via Central Limit Theorem, a Gaussian form.

In the remainder of this Section, we discuss a modification of the models proposed by Simon and Sutton. We show that if one changes the basic assumption of "equal assignment probabilities" of the business opportunities, the shape of the growth rates distribution is consequently modified and is no longer a Gaussian. We basically retain the island models approach and describe the growth of a firm as a two steps process. In the first step there is a random assignment among firms of a fixed number of business opportunities. The assignment procedures leads to a possible realization of the random variables $\mathbf{G}_i \quad \forall i \in \{1, \ldots, N\}$. In the second step, the $\mathbf{G}_i$ business opportunities assigned to firm i act as the source of micro-shocks affecting its size.

The first step of our model is based on a simple stochastic partition of a finite number of business opportunities, say M, among a population of N identical firms[7]. Instead of assuming, as common in the cited models, that the assignment of each opportunity is an independent event with constant probability $1/N$, we introduce the idea of ”competition among objects whose *market success*. . . [is] cumulative or self-reinforcing”(Arthur , 1994, 1996). We model this idea with a process where the probability for a given firm to obtain new opportunities depends on the number of opportunities already caught. Such a procedure of sequential assignment of M business opportunities among N firms is easily described using a Polya's urn scheme.

Consider an urn containing N balls of N different types. In this urn there is one ball for each type and each type represents a specific firm. A ball is drawn at random, then it is replaced and, moreover, 1 ball of the type drawn is added. Another random drawing is made from the ”new” urn containing one more ball and this procedure is repeated M times. It is straightforward to notice that in this way we introduce the desired effect that the drawing of either type increases the probability of the same type to be drawn at the next step (See Fig. 7 for a simple graphical exemplification of a first step of this procedure). We can now interpret the drawing of the ball of type $i \in \{1, \ldots, N\}$ as the assignment of one opportunity to firm i. In this context the outcome of each process of assignment of all the opportunities among firms is completely described by the occupancy N-tuple $(m_1, m_2, \ldots, m_N)$ where $\sum_{j=1}^{N} m_j = M$. The probability of obtaining a particular N-tuple (Feller (1968), p.120) is given by:

$$\mathbf{P}\left\{(m_1, m_2, \ldots, m_N)\right\} = \frac{1}{\binom{N+M-1}{N-1}} \quad . \tag{7}$$

The conditional probability of the same N-tuple, given that h opportunities have already been assigned to a single firm, can be derived from (7) simply noting that the problem remains exactly the same but the number of firms and the number of opportunities involved reduce to $N-1$ and $M-h$ respectively. Hence this conditional probability becomes:

$$\mathbf{P}\left\{(m_2, \ldots, m_N) \;\middle|\; m_1 = h\right\} = \frac{1}{\binom{N+M-h-2}{N-2}} \quad . \tag{8}$$

It is now easy to compute the marginal probability that a given firm obtains exactly h opportunities:

$$p(h; N, M) = \frac{\mathbf{P}\left\{(m_1, m_2, \ldots, m_N)\right\}}{\mathbf{P}\left\{(m_2, \ldots, m_N) \;\middle|\; m_1 = h\right\}} = \frac{\binom{N+M-h-2}{N-2}}{\binom{N+M-1}{N-1}} \tag{9}$$

which is the well known Bose-Einstein statistics[8]. To give an idea of the outcome of the previous assignment

[7]In accordance with the empirical investigations presented in the previous Section, we consider a fixed number of firms and abstract from any entry and exit dynamics.

[8]This statistics is mainly used in physics where it describes the peculiar thermodynamic behavior of a large family of subnuclear

procedure in Fig. 8 we compare the Bose-Einstein distribution with the Binomial distribution that would have been obtained if each opportunity were assigned with the same probability $1/N$. The "clustering" effect is evident in the fat tailed nature of the Bose-Einstein distribution, that suggests an increased probability to assign a large number of opportunities to a single firm. Furthermore, this distribution possesses a zero modal value in sharp contrast with the M/N value generated by an independent and equiprobable opportunities assignment.

The procedure just described provides a particular partition of M opportunities among N firms summarized by the N-tuple $(m_1, \ldots, m_N)$. As already mentioned, these business opportunities can be thought of as the source of micro-shocks affecting the size of the firm. We make no assumptions on the actual nature of these shocks and we want to relate "opportunities" to "growth" in the simplest way. Hence, we assume that these micro-shocks are randomly and independently drawn from a common distribution. Since we are interested only in the distribution of the relative growth rates, we assume that the shocks distribution has zero mean. The total growth of firm i is obtained adding $\mathbf{m}_i(t) + 1$ independent micro-shocks: $\mathbf{m}_i(t)$ shocks assigned by the Polya process plus the 1 already in the urn at the beginning of the assignment procedure. If $\mathbf{s}_i(t)$ stands for the random variable describing the (log) size of firm i at time t, the growth equation reads

$$\mathbf{s}_i(t+1) = \mathbf{s}_i(t) + \mathbf{g}_i(t) \qquad \mathbf{g}_i(t) = \sum_{j=1}^{\mathbf{m}_i(t)+1} \mathbf{x}_j(t)$$

where $\mathbf{x}$ are i.i.d. with a common distribution $F(x)$ with mean 0. Notice that the random growth rates $\mathbf{g}_i$ are identically, but not independently, distributed across firms, due to the global constraint $\sum_i \mathbf{m}_i = M$. Notice also that, being expressed in terms of growth rates, the effect of each opportunity on the size of the firm depends on the size itself. The unconditional probability distribution of $\mathbf{g}$, implied by the assignment procedure in (9), reads

$$F_{\mathrm{model}}(g; N, M, v_0) = \sum_{h=0}^{M} p(h; N, M)\ F(g)^{\bigstar(h+1)} \tag{10}$$

where $F(g)^{\bigstar h}$ stands for the h-time convolution of the micro-shocks distribution (i.e. the distribution of the sum of h micro shocks). The average number of opportunities per firm is M/N and if v_x is the variance of the micro-shock distribution, the distribution of growth rates $\mathbf{g}$ has mean 0 and variance $v = v_x(M/N+1)$.

At this point it is useful to clarify a few points about the assumptions just considered. First, concerning the zero mean hypothesis, notice that the choice of a distribution with a non-zero mean m_x would simply introduce an industry-wide growth trend proportional to Mm_x. In this paper we disregard this kind of trend since, in accordance with the empirical studies cited in Section 3, we describe the process in terms of market

particles.

shares. Second, both the hypotheses of no correlation among micro-shocks and of constant variance v_x in their distribution are working hypotheses introduced to keep the discussion clearer and can be relaxed, for instance introducing a mild correlation among micro-shocks or introducing a random variance extracted from a given distribution. Finally, in the next Section we will show that, when one considers a large number of firms N and a large average number of shocks per firm M/N, the actual choice for the shape of the micro-shocks distribution becomes irrelevant.

5 The source of the tent-shape

The mechanism presented in the previous Section is rather parsimonious in terms of the required parameters. It is able to provide a uniquely defined distribution for the firm growth rates when only three components are specified: the number of firms operating in the market N, the total number of "business opportunities" M representing the "sources" of the firms growth events and the effect that these events have on the size of the firm, captured by the micro-shocks probability distribution $F(x)$.

In this Section we analyze extensively the properties of the mechanism presented. Our aim is to understand under which conditions this mechanism is able to reproduce the empirical regularities described in Section 3. More precisely, we will show that when the number of firms N and the average number of micro-shocks per firm M/N become large, the growth rate distribution obtained from (10) progressively approaches a Laplace distribution.

In order to simplify the discussion, let us assume that the micro-shocks are normally distributed, with unit variance $v_x = 1$, i.e. $F(x) = N(x; 0, 1)$. This assumption is made only to keep the discussion easier, and we show in the next Section that our conclusions are largely independent from the actual shape of the micro-shock distribution.

We start our analysis with an example. Consider a sector with a reasonable number of firms, let say 100. This number is more or less of the same order of the population size in the manufacturing sectors analyzed in Section 3. Now suppose that no assignment of opportunities is performed, i.e. that $M = 0$. In other terms, each firm ends up with just one shock, the one originally put in the urn. Since the micro-shock distribution is $N(x; 0, 1)$, and exactly one shock is assigned to each firm, the observed growth rates distribution will have the same normal form. A picture of the associated density is reported in Fig. 9 with the label $M = 0$. The log scale on the y axis makes its parabolic shape clear. Now suppose instead to have a positive number of opportunities, for instance suppose that $M = 100$, so that the average number of opportunities per firm is now increased to 2. Now suppose that the micro shocks are distributed according to $N(x; 0, 1/2)$. This means that if the opportunities would be assigned independently, the firm growth rates would be the sum of two (the average number of shocks) normal variates with variance $1/2$, that is a normal variate with variance 1. Under the assumption of independent assignment, thus, the growth rates distribution would not change. But this is not our case. Indeed, if one follows the Polya process for the assignment of these 100 opportunities across the 100 firms, the growth rates density that emerges is different from a normal. This density is reported in Fig. 9 with the label $M = 100$ and is computed starting from the definition of the distribution in (10). As can be seen, the tails of the density are much fatter than in the normal case. This is a consequence of our mechanism of assignment: under the Polya process many opportunities tend to concentrate in few firms, producing final growth rates that are the sum of many micro-shocks and, consequently, are likely to become quite large. The shape of the density is, however, still close to a normal, at least in the central part. But what does happen if we further increase the number of opportunities? The answer is provided by the density shown in Fig. 9 with the label $M = 10000$. This is generated considering $M = 10000$ micro-shocks, i.e. an average of 101 micro-shocks per firm, distributed according to $N(x; 0, 1/101)$. Under the independent assignment hypothesis we would again obtain for the growth rates a normal density with unit variance. As can be seen, the obtained shape is instead almost identical to a tent-shaped Laplace distribution (see the inset of Fig. 9).

The agreement between the density generated by our assignment procedure and the Laplace can be further understood by looking at Fig. 10. Here we report the absolute deviation $|F_{\mathrm{model}}(g; M, N) - F_{\mathrm{L}}(g)|$ between the Laplace distribution and the distribution predicted by (10). We set the parameters to the same values used in Fig. 9. As can be seen, the absolute deviation is strongly reduced when the number of opportunities M is increased. On the other hand, its value seems to depend on g. In order to build a global measure of agreement between the two distributions that is independent from the value of g we consider the height of the maximum of the functions plotted in Fig. 10. We define

$$D(N, \lambda) = \sup_{-\infty < g < +\infty} |F_{\mathrm{model}}(g; N, M) - F_{\mathrm{L}}(g)| \tag{11}$$

where $F_{model}(g; N, M, f)$ is derived from the density defined in (10) and using normally distributed micro-shocks as described above, while $F_L(g)$ stands for the unit-variance Laplace distribution. The values of D for different N and M/N are plotted in Fig. 11. As N and M/N increase, the value of D decreases of several orders of magnitude and the ability of our mechanism to reproduce the Laplace distribution quickly improves. This picture tells us where, in the parameter space, we can expect that our mechanism gives a good account of the observed tent-like shape: this happens when both the number of firms N and the number of shocks per firm M/N are large. At this point a natural question arises: how large should this “large” be? Of course there are no definite answers to this question, since for any finite value of N and M the maximum absolute deviation of $F_{\rm model}$ from $F_{\rm L}$ is not zero. Indeed, in the next Section we will show that the perfect agreement can be reached only for asymptotically large values of these parameters. However, a quite satisfactory even if "heuristic" answer is provided by Fig. 12. Here, setting the value of N to "typical" values observed in data and choosing for M a sufficiently large value, we obtain the same level of agreement to the exponential shape found in empirical investigations. Notice that the binned density in Fig. 12 is computed using 7 independent realizations of the assignment process to provide direct comparability with the empirical plots in Section 3, where the 7 years of data were pooled together. We can conclude that the proposed mechanism of opportunities assignment, when the parameters N and M are set to proper (large) values and the micro-shocks are normally distributed, is able to reproduce the Laplace shape of the one-year growth rates density observed in industrial data.

We present now an analytical result proving that, as long as the total number of firms is large and their growth is generated by the assignment of a large number of small shocks, the double-exponential shape of the distribution of growth rates is robust to different specifications of the micro-shocks distribution. To show this, we study the model in the limit $M, N \to \infty$. In general, when we perform this limit, we cannot keep fixed the variance of the micro-shock distribution. Indeed, in order to match empirical observations, the growth rate distribution generated by the model must have finite variance[9]. If one increases the average number of micro-shocks assigned to firms, then the variance of these shocks must be decreased, in order to maintain the same variance for the final distribution, as we did in the example above. When the limit $M \to +\infty$ is considered, the rescaling of the variance becomes mandatory to avoid degenerate distributions. The mean of the Bose-Einstein distribution in (9) is M/N and, if the micro-shock distribution possesses finite variance v_x, the variance of the distribution $F_{\rm model}$ in (10) becomes $v_x\,(1 + M/N)$. In order to obtain a final distribution with finite variance v, when we vary the values of M and N, we must replace the random variable $\mathbf{x}$ in (10) with $\sqrt{vN}/\sqrt{v_x(N+M)}\,\mathbf{x}$. This is equivalent to consider a micro-shocks distribution of the form $F(\sqrt{(N+M)/(vN)}\,x)$ where $F(x)$ is a distribution with unit variance[10].

The main theoretical result of this paper is established in the following theorem that provides a complete asymptotic characterization of the distribution of firms growth rates implied by our model.

Theorem 1. Let $F(x)$ be a probability distribution with zero mean and unit variance and let $\mathbf{m}$ be a discrete

[9]More precisely, the empirical growth rate densities seem to be characterized by an asymptotic exponential behavior, thus possessing all the central moments. As the discussion below will however reveal, it is enough to assume, in all the cases of interest, the existence of the second moment of the micro-shocks distribution.

[10]We take an unit variance unscaled distribution F to get rid of the parameter v_x. This choice obviously does not constitute a reduction in generality, since any random variable with finite variance is proportional to a random variable with unit variance.

random variable distributed according to the Bose-Einstein distribution with parameters N and M, as defined in (9).

Consider the random variable

$$\mathbf{g} = \sum_{j=1}^{\mathbf{m}+1} \mathbf{x}_j \tag{12}$$

where $\mathbf{x}$ are i.i.d random variables distributed according to $F(\sqrt{(N+M)/(vN)}\, x)$.

When M and N go to infinity, if the limit of M/N exists, finite or infinite, the random variable g converges in distribution to a proper random variable whose specific distribution depends on the asymptotic order relation between N and M. If $F_{\text{model}}(g)$ is the distribution of $\mathbf{g}$ one has

1. If $M, N \to \infty$ and N asymptotically dominates M, that is $\lim_{M,N\to\infty} M/N = 0$, then

$$\lim_{M,N\to+\infty} F_{\text{model}}(g) = F(\frac{g}{\sqrt{v}}) \tag{13}$$

   i.e. the random variable $\mathbf{g}$ converges in distribution to the rescaled micro-shock random variable $\sqrt{v}\mathbf{x}$.

2. If $M, N \to \infty$ and N is asymptotically equivalent to M, i.e. $\lim_{M,N\to\infty} M/N = \lambda \in R^+$ then

$$\lim_{M,N\to\infty} F_{\text{model}}(g) = \Lambda(g; \lambda) \tag{14}$$

   where $\Lambda(g; \lambda)$ is a distribution function whose expression depends on λ and on the micro-shock distribution $F(x)$ and whose characteristic function can be completely specified in terms of the micro-shocks characteristic function.

3. If $M, N \to \infty$ and M asymptotically dominates N, that is $\lim_{M,N\to\infty} N/M = 0$, then

$$\lim_{M,N\to\infty} F_{\text{model}}(g) = F_L(g; 0, \sqrt{v/2}) \tag{15}$$

   where $F_L(g; 0, \sqrt{v/2})$ is a Laplace distribution with mean 0 and variance v.

Proof. See Appendix A.

The discussion of the three possible cases in Theorem 1 can help to clarify the basic intuition behind it.

In the first case, the number of firms N grows faster than the number of assigned shocks M. The probability of a firm to end up with the single shock he had from the beginning tends asymptotically to one, so that the growth rate distribution converges to the distribution of this single shock.

In the second case, the average number of shocks per firm M/N tends to a constant value λ and the growth rate tends to a distribution which is a "distorted" version of the micro-shock distribution F, the "distortion" being the effect of the random assignment of shocks.

Finally, in the third case, the number of micro-shocks assigned to firms increases more rapidly that the number of firms, so that the number of shock a firm gets becomes, on average, infinite. At the same time, however, the variance of each shock, which is proportional to $N/(N+M)$, decreases toward zero. The mixed effect of an infinite number of shock of infinitesimal magnitude leads to a phenomenon similar to the central limit theorem: the limit distribution is independent from the exact specification of the micro-shock distribution and displays an "universal" character. In this case, instead of the Gaussian distribution implied by the standard Central Limit Theorem, we obtain a Laplace distribution.

Before ending this Section let us briefly discuss the second evidence highlighted in Section 3, namely the progressive "normalization" of the growth rates density when one considers longer time scales (see Fig. 5). If one assumes that the process of opportunities assignment is repeated anew each year, i.e. that no memory of the previous year assignment is retained when the new year opportunities are assigned, then the growth rates of each year are independent for any firm. Consequently, the T lags growth rates are the sum of T independent random variables and, when T becomes large, their distribution tends toward a Gaussian. In this way we recover, at least as a first approximation, the behavior reported in Fig. 6. One can however argue that the idea of introducing strong positive-feedback effects in the opportunities assignment inside the same year and no memory at all of the previous year assignment sounds rather inconsistent. After all, if dynamic increasing returns are there, why should they disappear during the new year's eve? We believe that the relevant point to notice here is that the one-year time span used to build empirical databases does not posses any meaning inside our model and, most probably, even inside the real economic dynamics (c.f. the discussion in Geroski (2000)). In this respect, one can think that the assignment procedure of our model works on a certain time span, let say on a time scale from 6 to 36 months, but that for longer time period the effect of the past captured opportunities fades away. This reduction in the relevance of opportunities caught in the far past can be progressive and smooth. From evidence shown in Fig. 6, we can suppose that the reduction becomes relevant on a time scale of few years and, plausibly, acts with a different strength in different sectors. In order to describe this kind of dynamics one can modify the assignment mechanism introduced in Section 4 assuming, for instance, that the balls of a given firm are removed from the urn after a given time span or that their contribution to the probability of capturing new balls is inversely proportional to their "age", i.e. the number of turns they stayed in the urn. This kind of models would consider explicitly the flow of time and, consequently, introduce quite a few technical difficulties. We do not want to pursue here this issue but it is clear that our model should not be considered valid on a very long time scale[11].

6 Conclusions

This paper presents crucial evidence in support of the tent-shape of the firm growth rates distribution, extending previous findings in two different directions. First, we replicate the analysis already performed by Stanley et al. (1996) and Bottazzi et al. (2001) on a new databank (MICRO.1) covering many firms of the Italian manufacturing industry. Second, using data disaggregated by sector, we prove that the shape of these distributions is not a mere effect of aggregation. Although intersectoral differences clearly arise, we conclude, in line with previous studies, that the tent-shape (double exponential) distribution of corporate growth rates appears as an extremely robust feature of the manufacturing industry, characterized by a higher regularity than the one shown by size distributions.

On the theoretical side, we propose a model which describes the dynamics of firms growth. The model clearly originates in the Simon inspired literature on firm dynamics with which it shares two central features. First, different firms are viewed as different realizations of the same stochastic process. Second, the model includes a very simple idea of competition represented by a global constraint on the total number of available growth opportunities.

The essential novelty of our approach lies in the assignment procedure of different business opportunities among different firms. In our model, the probability for a given firm to obtain new opportunities depends on the number of opportunities already caught. In this way, we introduce dynamic increasing returns in the growth process of firms. Economies of scale, economies of scope, network externalities and knowledge accumulation are just a few examples of possible economic mechanisms able to generate positive feedbacks within markets, businesses and industries. The overall effect can be described as the emergence of a sort of ”attracting force” between the various opportunities that tends to group them in bigger chunks leading to the appearance of two noticeable properties in their unconditional distribution: the presence of a fat tail, which indicates a more likely presence of extremely large number of opportunities assigned to a single firm, and the absence of a natural scale of the underlying process, hinted by the 0 value of the mode.

The ability of the model to reproduce empirical findings without requiring a fine tuning of the parameters is due to the Theorem in Section 5 and constitutes its main strength. This theorem ensures that, when the

[11] In fact, the present model would generate, in the long run, log-normal size distributions that are not observed in empirical analyses. Concerning the MICRO.1 database considered in Section 3, the shapes of the size distributions in the different sectors appear quite heterogeneous. See Bottazzi et al. (2003) for details.

number of firms and the number of opportunities per firm go to infinity, the growth rates distribution generated by the model converges to the Laplace. According to this Theorem, the sole requirement is that the number of “business opportunities” for which firms compete is increasingly larger than the number of competing firms. Consequenlty, the competitive success is not seen as the outcome of a single lucky event granting one firm a persistent, dominant, position but, rather, as the ability of a firm to build its new success, through a permanent struggle and inside an extremely volatile environment, on the basis of its past, sucessfull, behaviour. If this requirement is fullfilled, neither the fine tuning of the parameters values nor the choice of a particular micro-shocks distribution are required for our model to reproduce the observed tent-shape of the distribution of growth rates.

The model presented here can be extended to capture also the scaling relationship between the variance of the growth rates and the size of the firm discussed in Stanley et al. (1996). This extension can be easily performed considering a diversified firm competing in independent sub-sectors whose number depend on the firm size. We did not pursue this issue here since the empirical evidence on this point seems mixed: the relation between growth rates variance and firm size is present in the COMPUSTAT database both at aggregate (Stanley et al., 1996) and disaggregated (Bottazzi and Secchi, 2004) level and in the worldwide pharmaceutical industry (Bottazzi et al., 2001) but, as mentioned, it seems to be absent in the Italian manufacturing industry.

We are aware that the Polya urn mechanism presented in Section 4 does only constitute a simple metaphor of “positive feedback that operates - within markets, businesses, and industries - to reinforce that which gain success or aggravate that which suffer loss” (Arthur (1994), p.100). The direct test of this assumption, dealing with the intimate essence of the competitive dynamics, is not trivial since it would require the joint investigation of the existence and nature of increasing returns both at the level of single firm and of the whole industry and, consequently, would not allow an explicit identification of a single hypothesis.

However, one can think to the body of empirical literature on the “clusterization” of technological innovations (Silverberg and Verspagen, 2000), the increasing returns in research activity (Henderson and Cockburn, 1996) and the self-reinforcing effect in the creation of managerial talents (Penrose, 1958) as suggestive evidences supporting the existence of an underlying positive feedback mechanism shaping the competitive dynamics.

From a more general point of view, the justification of the stochastic models like the one presented here lies in their ability to describe the observed regularities. At the same time, their relative value should be expressed in terms of their degree of generality and by their ability of “explaining why the generalization ’should’ fit the facts” (Simon, 1968).

A Asymptotic behavior of the growth rates density

In this Appendix we prove Theorem 1 concerning the asymptotic behavior of firms growth rates distribution (10). Before attacking the main proof, we need two preliminary results on the asymptotic properties of the Polya process. To this purpose, it is convenient to introduce the Bose-Einstein generating function, formally defined as

$$Q(z;N,M)=\sum_{h=0}^{M} z^h \; p(h;N,M) \quad . \tag{A1}$$

A more suitable representation of this function is provided by the following

Lemma 1. The generating function in (A1) admits the following integral representation

$$Q(z;N,M)=(N-1)\int_0^1 dt\;(1-t)^{N-2}\;(1-t+t\,z)^M \tag{A2}$$

Proof. Consider the generic term in (9), expanding the binomial coefficient one obtains

$$p(h;N,M)=\frac{\Gamma(N+M-h-1)}{\Gamma(N-1)\;\Gamma(M-h+1)}\;\frac{\Gamma(N)\;\Gamma(M+1)}{\Gamma(N+M)} \tag{A3}$$

where we used the relation $a!=\Gamma(a+1)$. Multiplying both the numerator and denominator by $\Gamma(h+1)$ and using the definition of the beta function $B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$ one can rewrite (A3) as

$$p(h;N,M)=(N-1)\binom{M}{h}\;B(N+M-h-1,h+1) \tag{A4}$$

where the second factor stands, with usual notation, for the binomial coefficient. Using the following integral representation of the beta function (Abramowitz and Stegun (1966) p.258)

$$B(a,b)=\int_0^1 dt\;t^{a-1}\;(1-t)^{b-1} \tag{A5}$$

and substituting it in (A4) one obtains

$$p(h;N,M)=(N-1)\binom{M}{h}\int_0^1 dt\;t^h\;(1-t)^{N+M-h-2} \quad . \tag{A6}$$

The integral representation (A6) can be substituted for each term in (A1) to obtain

$$Q(z;M,N) = (N-1) \sum_{h=0}^{M} \binom{M}{h} z^h \int_0^1 dt\, t^h\, (1-t)^{N+M-h-2} \tag{A7}$$

that, once the summation on h is moved inside the integral and the binomial expansion is collected, reduces to (A2). *Q.E.D.*

Now formally rewrite (A2) as

$$Q(z;N,M) = -\int_0^1 dt\, (1-t+t\,z)^M \frac{d}{dt} (1-t)^{N-1} \tag{A8}$$

so that, after a straightforward integration by parts, one obtains

$$Q(z;N,M) = 1 - (1-z)\, \frac{M}{N}\, Q(z;N+1,M-1) \quad . \tag{A9}$$

This recurrence relation can be used to characterize the asymptotic behavior of the Bose-Einstein generating function Q for large values of the parameters M and N. The result is provided by the following

Lemma 2. When $|z| \leq 1$ if the values of N and M tend to infinity, the generating function $Q(z;N,M)$ tends to the fixed point of the recurrence relation (A9). More precisely

$$\lim_{M,N\to\infty} Q(z;N,M) \sim \frac{1-(1-z)\,\frac{M}{N^2}\, E(N,M)}{1+(1-z)\,\frac{M}{N}} \qquad |z| \leq 1 \tag{A10}$$

where E is a bounded function of M and N, that is $\lim_{M,N\to\infty} |E| < \infty$.

Proof. Let us consider the difference between $Q(z;N,M)$ and $Q(z;N+1,M-1)$. Using the integral representation in (A2) it can be written as

$$\begin{aligned} &Q(z;N,M) - Q(z;N+1,M-1) = \\ &\int_0^1 dt (1-t)^{N-2}\, (1-t+zt)^{M-1}\, \left[(N-1)\,t\,z - (1-t)\right] \quad . \end{aligned} \tag{A11}$$

Splitting the square brackets and using again the representation (A2) one obtains

$$\begin{aligned} &Q(z;N,M) - Q(z;N+1,M-1) = \\ &z\,(N-1) \int_0^1 dt\, t\, (1-t)^{N-2}\, (1-t+zt)^{M-1} - \frac{1}{N}\, Q(z;N+1,M-1) \quad . \end{aligned} \tag{A12}$$

Since $|a-b| \le |a| + |b|$, the absolute value of the left-hand side of (A12) satisfies

$$|Q(z;N,M) - Q(z;N+1,M-1)| \le \\ |z|\,(N-1)\,\int_0^1 dt\; t\,(1-t)^{N-2}\;|1-t+zt|^{M-1} + \frac{1}{N}\;|Q(z;N+1,M-1)| \quad . \qquad \text{(A13)}$$

If $|z| \le 1$ it is $|1-t+zt| \le 1$ $\forall t \in [0,1]$. Then, the value of the first integral on the right hand side of (A13) does not decrease if one replaces $|1-t+zt|$ with 1 so that, after an integration by parts, one obtains

$$|Q(z;N,M) - Q(z;N+1,M-1)| \le \frac{|z|}{N} + \frac{1}{N}\;|Q(z;N+1,M-1)| \quad . \qquad \text{(A14)}$$

From the expansion in (A1) immediately follows that when $|z| \le 1$ it is $|Q(z;N,M)| \le Q(1;M,N) = 1$ so that

$$|Q(z;N,M) - Q(z;N+1,M-1)| \le \frac{1+|z|}{N} \le \frac{2}{N} \qquad \text{(A15)}$$

and one can write

$$Q(z;N+1,M-1) = Q(z;N,M) + \frac{1}{N}E(N,M) \qquad \text{(A16)}$$

where E is a bounded function of N and M. Substituting (A16) in (A9) proves the assertion. *Q.E.D.*

We can now use the result of the previous Lemma to proof Theorem 1.

Proof of Theorem 1. From (12) the model growth rates distribution function reads

$$F_{\text{model}}(g) = \sum_{h=0}^{M} p(h;N,M)\; F(\sqrt{(N+M)/(vN)}x)^{\star(h+1)} \qquad \text{(A17)}$$

and, since the series in (A17) is absolutely convergent, its characteristic function satisfies

$$\phi_{\text{model}}(k) = \int_{-\infty}^{+\infty} e^{i\,k\,g}\; dF_{\text{model}}(g;N,M) = \sum_{h=0}^{M} p(h;N,M)\;\phi(\sqrt{vN/(N+M)}\,k)^{h+1} \qquad \text{(A18)}$$

where $\phi(k)$ is the characteristic function of the micro-shock distribution $F(x)$ and we used the fact that the characteristic function of the h-convolution of the distribution F is h times its characteristic function.

Remembering the definition of the Bose-Einstein generating function in (A1) one can rewrite the characteristic function (A18) as

$$\phi_{\text{model}}(k;N,M) = \phi(\sqrt{vN/(N+M)}\,k)\; Q(\phi(\sqrt{vN/(N+M)}\,k);N,M) \quad . \qquad \text{(A19)}$$

This expression clearly separates the contribution of the random procedure used to assign the growth shocks to the firms, which is responsible for the term Q, from the specific nature of the micro-shocks distribution, which is described by the characteristic function ϕ.

Using the relation in (A19) one can express the asymptotic behavior of the characteristic function of the model in terms of the asymptotic behavior of the generating function.

Since $\phi(k)$ is a characteristic function it is $|\phi(k)| \leq 1, \forall k$ (Lemma 1, Feller (1968) p.499) and one can use (A10) to obtain

$$\phi_{\text{model}}(k) \sim \phi(\sqrt{(M+N)/vN}\,k)\frac{1-(1-\phi(\sqrt{vN/(N+M)}\,k))\,\frac{M}{N^2}\,E(N,M)}{1+(1-\phi(\sqrt{vN/(N+M)}\,k))\,\frac{M}{N}} \tag{A20}$$

for large values of M and N. Now consider in turn the three possibilities above:

1. if M and N go to infinity in such a way that N diverges faster than M and $\lim_{M,N\to\infty}\ M/N = 0$, the rescaled micro-shock characteristic function $\phi(\sqrt{vN/(N+M)}\,k)$ tends to $\phi(\sqrt{v}\,k)$ and $Q(z;N,M)$ tends to 1 so that

$$\lim_{M,N\to+\infty} \phi_{\text{model}}(k) = \phi(\sqrt{v}\,k) \tag{A21}$$

2. the second case is analogous to the first. If M and N are of the same asymptotic order and $\lim_{M,N\to\infty}\ M/N = \lambda$, obviously it is $\lim_{M,N\to\infty}\ M/N^2 \to 0$ and taking the limit of (A20) one has

$$\lim_{M,N\to+\infty} \phi_{\text{model}}(k) = \frac{\phi(\sqrt{v/(1+\lambda)}\,k)}{1+\lambda-\lambda\phi(\sqrt{v/(1+\lambda)}\,k)} \tag{A22}$$

3. if M diverges faster then N so that $\lim_{M,N\to\infty}\ N/M = 0$, the micro-shock variance goes to zero proportionally to N/M. Since the micro-shocks distribution $F(x)$ possesses unitary second moment it admits the following expansion around the origin (Lemma 2, Feller (1971) p. 512)

$$\phi(\sqrt{vN/(N+M)}\,k) = 1 - \frac{1}{2}\,k^2\,\frac{vN}{M+N} + o\left(\frac{vN}{M+N}\right) \quad . \tag{A23}$$

Substituting this expansion in (A20) and taking the limit $N/M \to 0$ one obtains

$$\lim_{M,N\to+\infty} \phi_{\text{model}}(k) = \frac{1}{1+\frac{1}{2}\,v\,k^2} \tag{A24}$$

where the left-hand side is the characteristic function of a Laplace distribution with mean 0 and variance v, as can be easily checked computing the Fourier transform of (1).

The limits in (A21), (A22) and (A24) are pointwise limits involving the characteristic function of the model. Since $F(x)$ possesses a finite variance, the function $\phi(k)$ is continuous in the origin and so are all the limit functions in the above equations.

Since the point-like limit of a sequence of characteristic functions to a function continuous in the origin implies the limit in distribution for the associated sequence of random variables (Theorem 2, Feller (1971) p. 508), the assertion is proved. *Q.E.D.*

References

Abramowitz, M. and Stegun, I.A. *Handbook of Mathematical Functions*. New York: Dover, 1964.

Agrò, G. "Maximum Likelihood Estimation for the Exponential Power Function Parameters." *Communication in Statistics - Simulation and Computation*, Vol. 24(1995), pp. 523-536.

Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, F., Salinger, M.A., Stanley, H.E. and Stanley, M.H.R. "Scaling Behavior in Economics: I. Empirical Results for Company Growth." *Journal de Physique I France*, Vol. 7(1997), pp. 621-633.

Amaral, L.A.N., Gopikrishnan, P., Plerou, V. and Stanley, H.E. "A Model for the Growth Dynamics of Economic Organizations." *Physica A*, Vol. 299(2001), pp. 127-136.

Arthur, B. W. *Increasing Returns and Path Dependence in the Economy*. Ann Arbor: University of Michigan Press, 1994.

Arthur, B. W. "Increasing Returns and the New World of Business." *Harvard Business Review*, July-August(1996), pp. 100-109.

Bottazzi, G. and Secchi A. "Common Properties and Sectoral Specificities in the Dynamics of U.S. Manufacturing Companies." *Review of Industrial Organization*, Vol. 23(2004), pp. 217-232.

Bottazzi, G., Cefis E., Dosi G. and Secchi, A. "Invariances and Diversities in the Evolution of Manufacturing Industries." L.E.M. Working Paper S.Anna School of Advanced Studies no. 2003-21, 2003.

Bottazzi, G. and Secchi, A. "Why are Distributions of Firm Growth Rates Tent-shaped?" *Economics Letters*, Vol. 80(2003), pp. 415-420.

Bottazzi, G., Cefis, E., Dosi, G. "Corporate Growth and Industrial Structure. Some Evidence from the Italian Manufacturing Industry." *Industrial and Corporate Change*, Vol. 11(2002), pp. 705-723.

Bottazzi, G., Dosi, G., Lippi, M., Pammolli, F. and Riccaboni, M. “Innovation and Corporate Growth in the Evolution of the Drug Industry.” *International Journal of Industrial Organization*, Vol. 19(2001), pp. 1161-1187.

Dunne, T., Roberts, M.J. and Samuelson, L. “The Growth and Failure of U.S. Manufacturing Plants.” *Quarterly Journal of Economics*, Vol. 104(1988), pp. 671-698.

Ericson, R. and Pakes, A. “Markov-Perfect Industry Dynamics: a Framework for Empirical Work.” *Review of Economic Studies*, Vol. 62(1995), pp. 53-82.

Evans, D.S. “The Relationship between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries.” *Journal of Industrial Economics*, Vol. 35(1987), pp. 567-581.

Evans, D.S. “Test of Alternative Theories of Firm Growth.” *Journal of Political Economy*, Vol. 95(1987), pp. 657-674.

Feller, W. *An Introduction to Probability Theory and Its Applications Vol.I*. New York: Wiley & Sons, 1968.

Feller, W. *An Introduction to Probability Theory and Its Applications Vol.II*. New York: Wiley & Sons, 1971.

Geroski, P.A. “Growth of Firms in Theory and in Practice.” In Foss N. and Mahnke V., eds., *New Directions in Economics Strategy Research*. Oxford: Oxford University Press, 2000.

Gibrat, R. *Les Inégalités Économiques*. Paris: Librairie du Recueil Sirey, 1931.

Hall, B. H. “The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector.” *Journal of Industrial Economics*, Vol. 35(1987), pp. 583-606.

Hart, P.E. and Prais, S.J. “The Analysis of Business Concentration.” *Journal of the Royal Statistical Society*, Vol. 119(1956), pp. 150-191.

Henderson, R. and Cockburn, I. “Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery.” *The RAND Journal of Economics*, Vol,. 27(1996), pp. 32-59.

Hymer, S. and Pashigian, P. “Firm Size and Rate of Growth.” *Journal of Political Economy*, Vol. 70(1962), pp. 556-569.

Ijiri, Y. and Simon, H.A. *Skew Distributions and the Sizes of Business Firms*. Amsterdam: North Holland Publishing Company, 1977.

Jovanovic, B. “Selection and the Evolution of Industry.” *Econometrica*, Vol. 50(1982), pp. 649-70.

Kalecki, M. “On the Gibrat Distribution.” *Econometrica*, Vol. 13(1945), pp. 161-170.

Lotti, F., Santarelli, E. and Vivarelli, M. “Does Gibrat’s Law Hold Among Young, Small Firms?” *Journal of Evolutionary Economics*, Vol. 13(2003), pp. 213-235.

Nelson, R.R. and Winter, S.G. “Forces Generating and Limiting Concentration under Schumpeterian Competition.” *The Bell Journal of Economics*, Vol. 9(1978), pp. 524-48.

Pakes, A. and Ericson, R. “Empirical Implications of Alternative Models of Firm Dynamics.”, *Journal of Economic Theory*, Vol. 79(1988), pp. 1-45.

Penrose, E. *The Theory of the Growth of the Firm*. Oxford: Oxford University Press, 1958.

Stanley, M.H.R., Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M. A. and Stanley, H.E. “Scaling Behavior in the Growth of Companies.” *Nature*, Vol. 379(1996), pp. 804-806.

Silverberg, J. and Verspagen, B. “Breaking the Waves: a Poisson Regression Approach to Schumpeterian Clustering of Basic Innovations.” Eindhoven Centre for Innovation Studies Ecis W.P. 00.16, 2000.

Simon, H.A. “On Judging the Plausibility of Theories.” In B. Van Rootselaar and J.F. Staal, eds., *Logic, Methodology and Philosophy of Sciences Vol.III*. Amsterdam: North Holland, 1968.

Simon, H.A. and Bonini, C.P. “The Size Distribution of Business Firms.” *American Economic Review*, Vol. 48(1958), pp. 607-617.

Singh, A. and Whittington, G. “The Size and Growth of Firms”, *Review of Economic Studies*, Vol. 42(1975), pp. 15-26.

Steindl, J. *Random Processes and the Growth of Firms*. London: Griffin, 1965.

Subbotin, M.T. “On the Law of Frequency of Errors.” *Matematicheskii Sbornik*, Vol. 31(1923), pp. 296-301.

Sutton, J. “Gibrat’s Legacy” *Journal of Economic Literature*, Vol. 35(1997), pp. 40-59.

Sutton, J. *Technology and Market Structure, Theory and History*. Cambridge, Mass.: MIT Press, 1998.

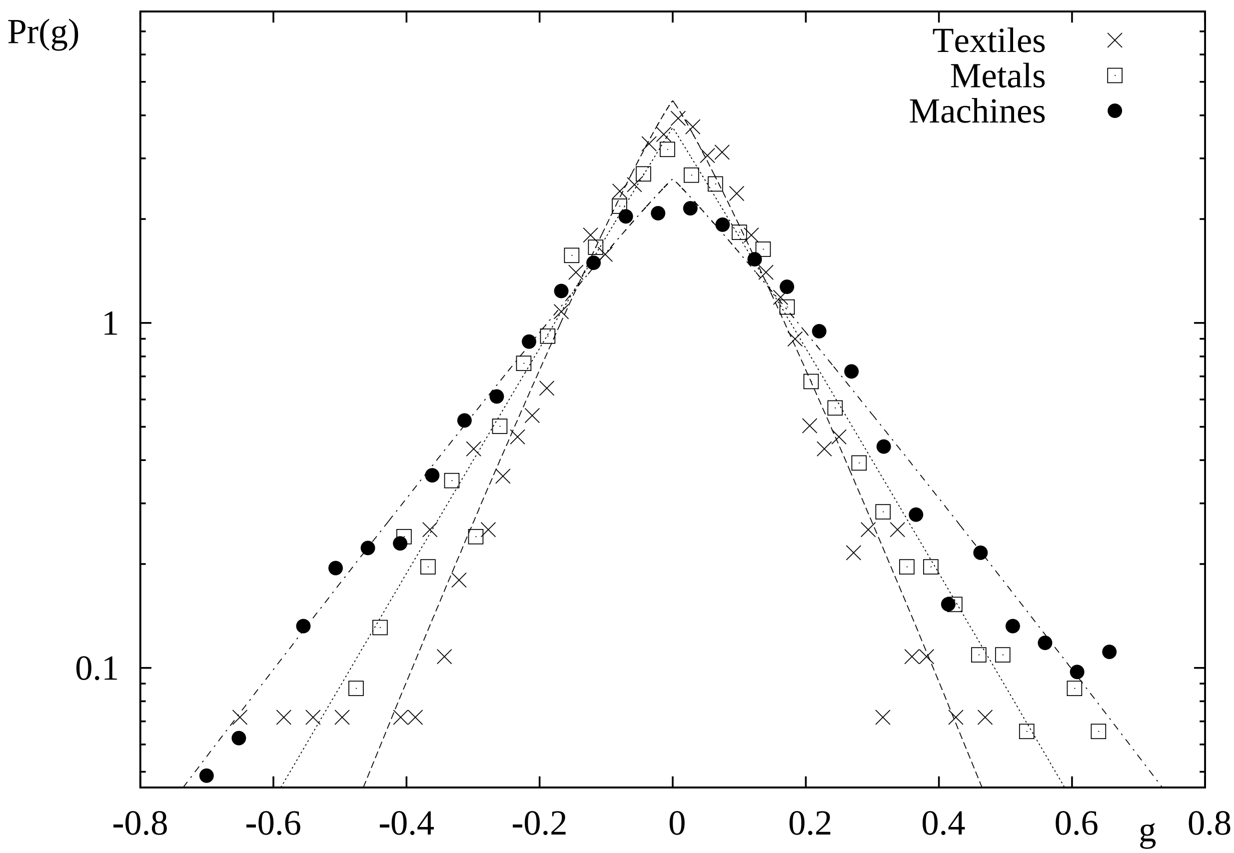

Figure 1: Binned empirical densities of the growth rates for the three sectors of textiles finishing, treatment and coating of metals and special purpose (metallurgy, mining, chemistry ...) machinery. The densities are pooled over all the 7 years (Data source: MICRO.1). Notice the log scale on the y-axes.

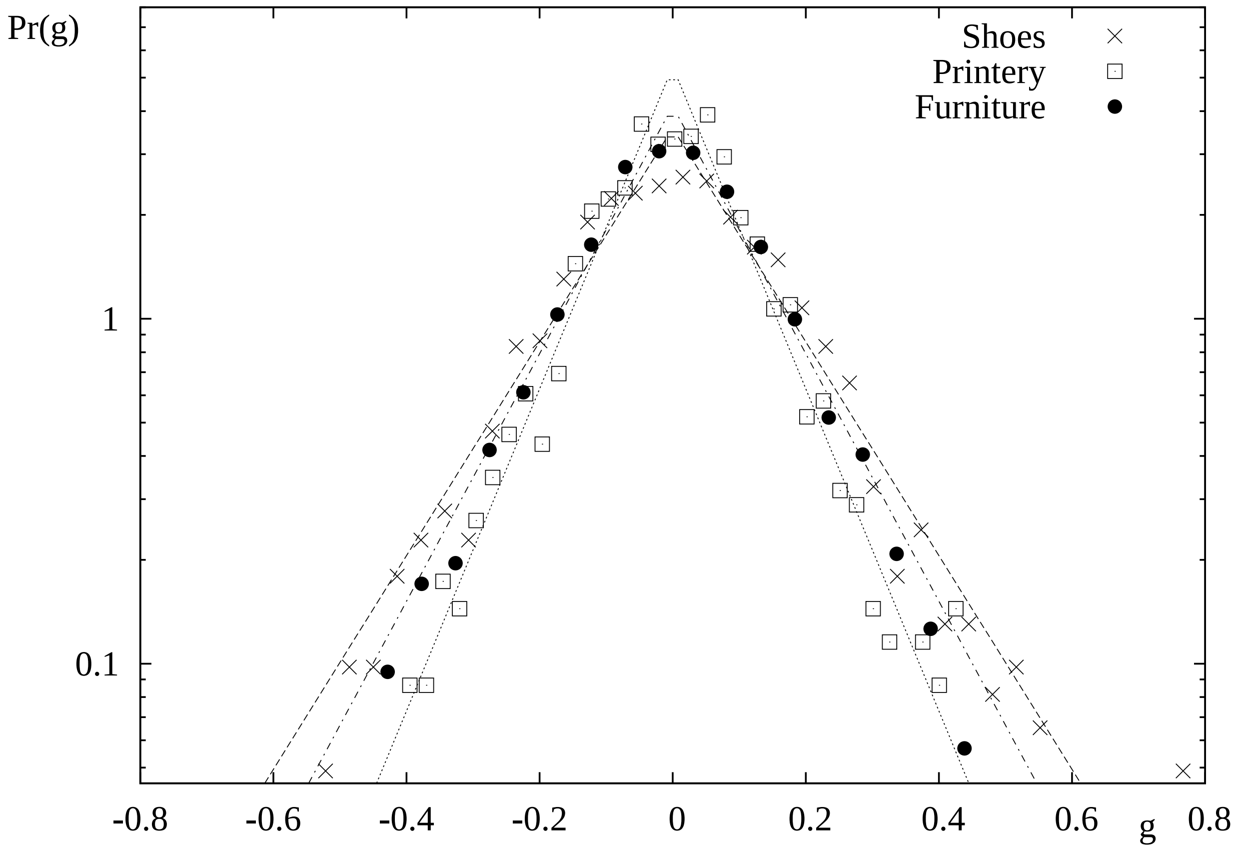

Figure 2: Binned empirical densities of the growth rates for the three sectors of footwear, printing and furniture. The densities are pooled over all the 7 years (Data source: MICRO.1). Notice the log scale on the y-axes.

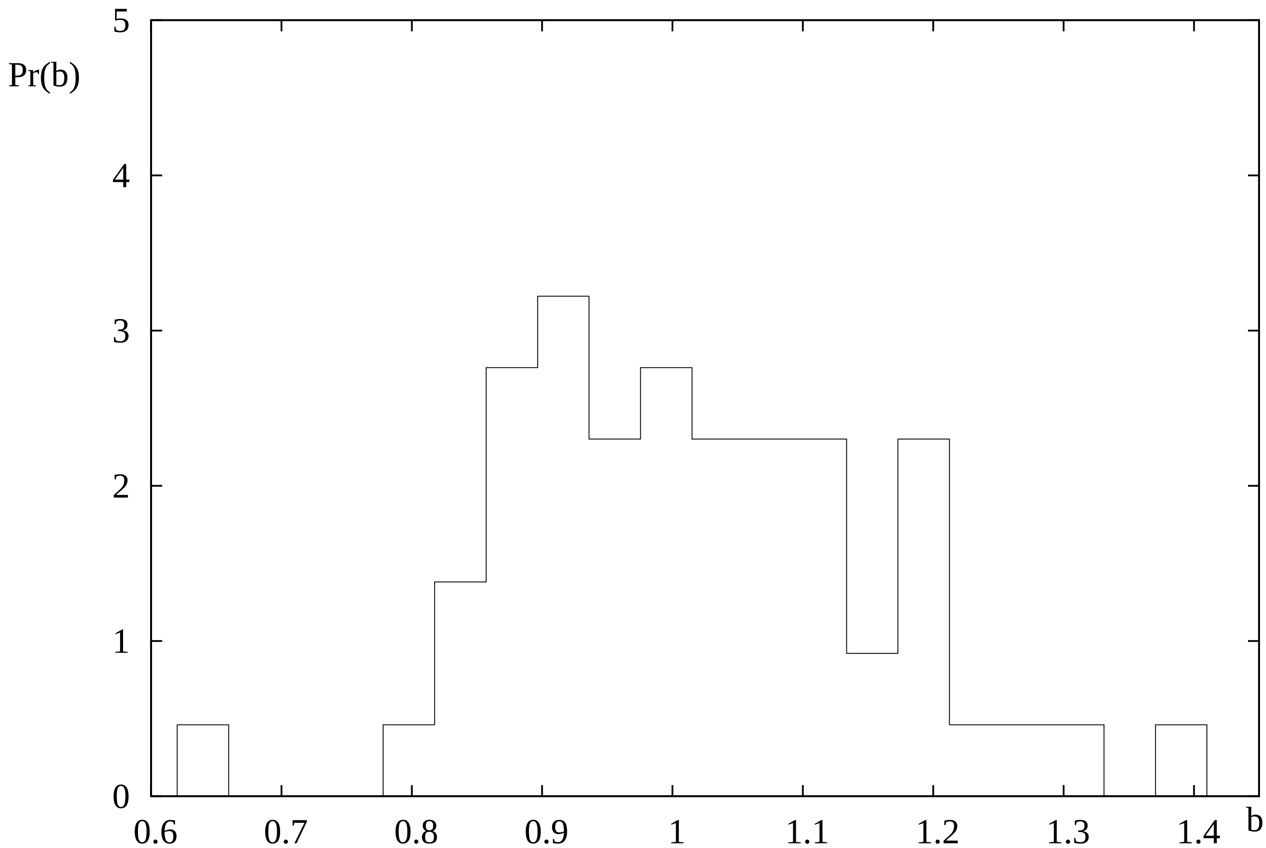

Figure 3: The binned empirical density of the b parameter values estimated using maximum likelihood over the different sectors. The values for the different sectors together with standard errors are reported in Table 1.

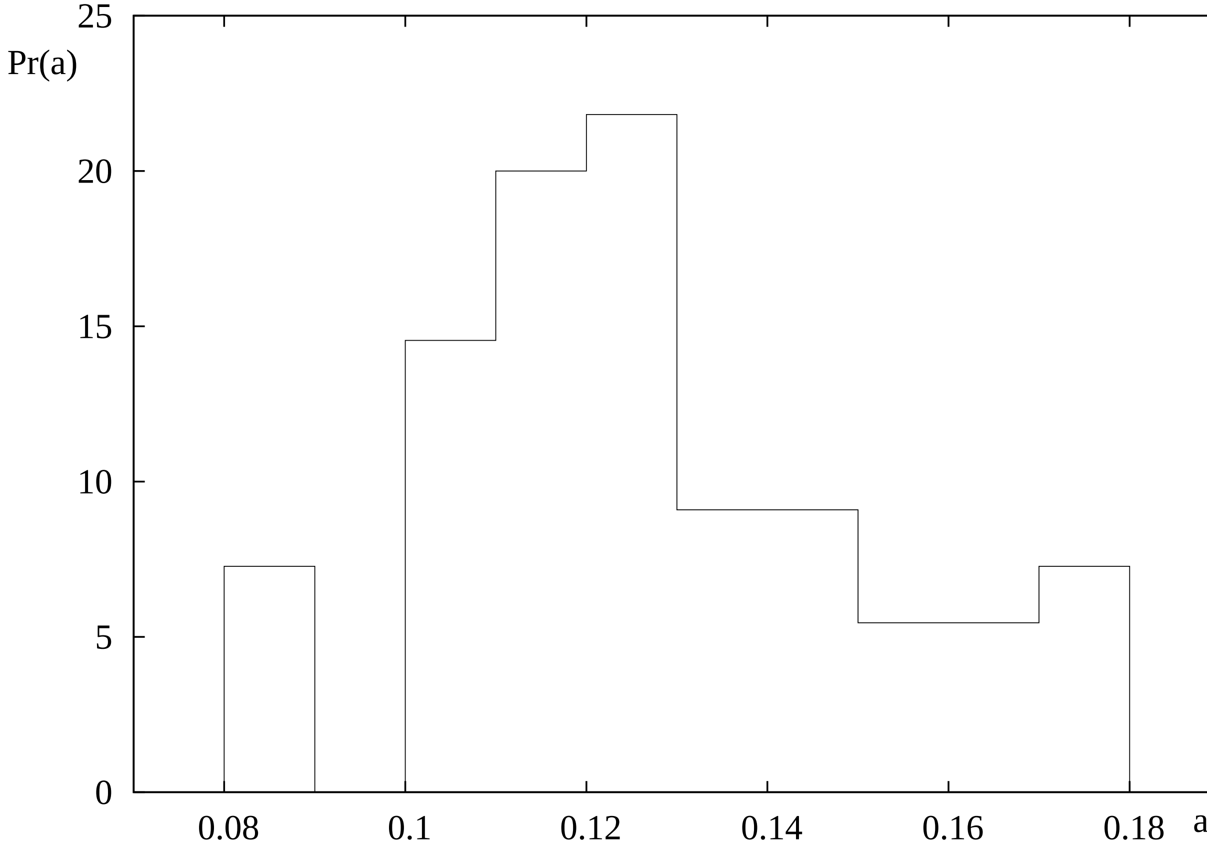

Figure 4: The binned empirical density of the a parameter values estimated using maximum likelihood over the different sectors. The values for the different sectors together with standard errors are reported in Table 1.

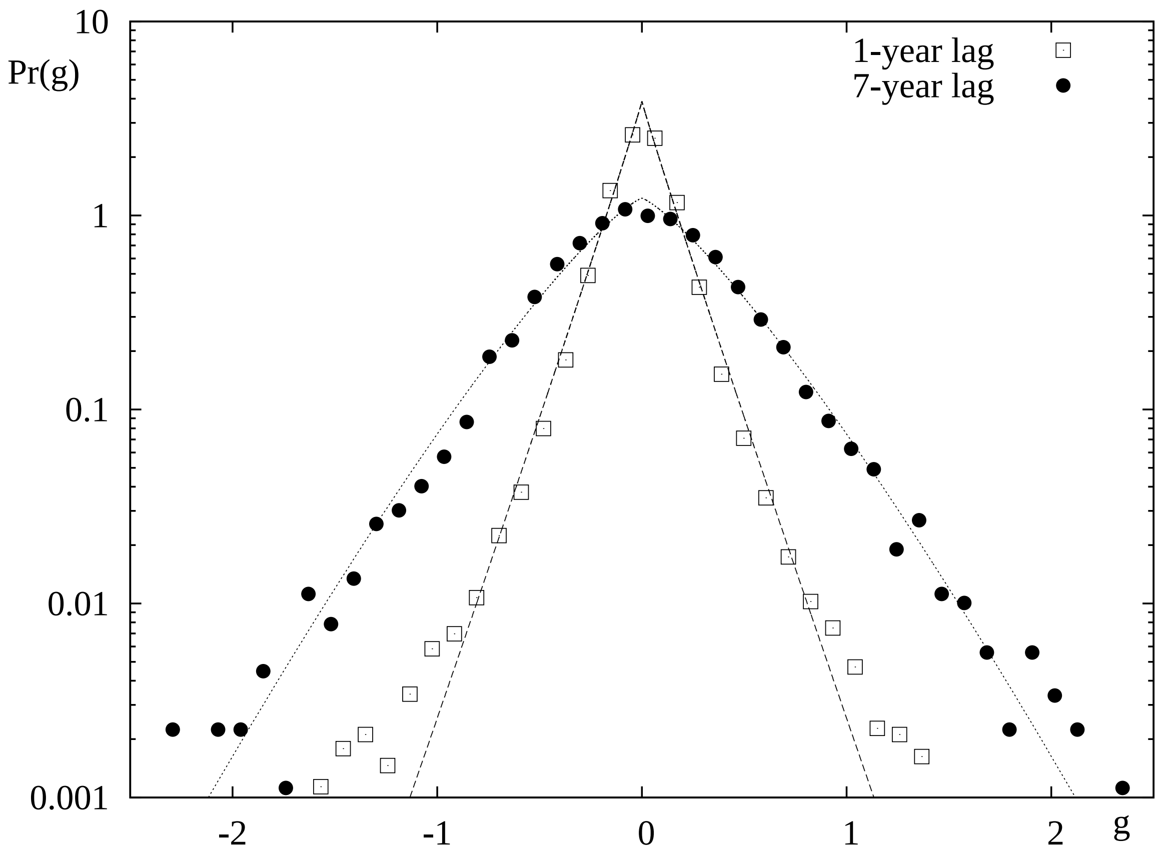

Figure 5: Binned empirical densities of the growth rates of the whole Italian manufacturing industry (Data source: MICRO.1) together with the best Subbotin fit. Two different time lags are considered of 1 and 7 years respectively. Notice the log scale on the y-axes.

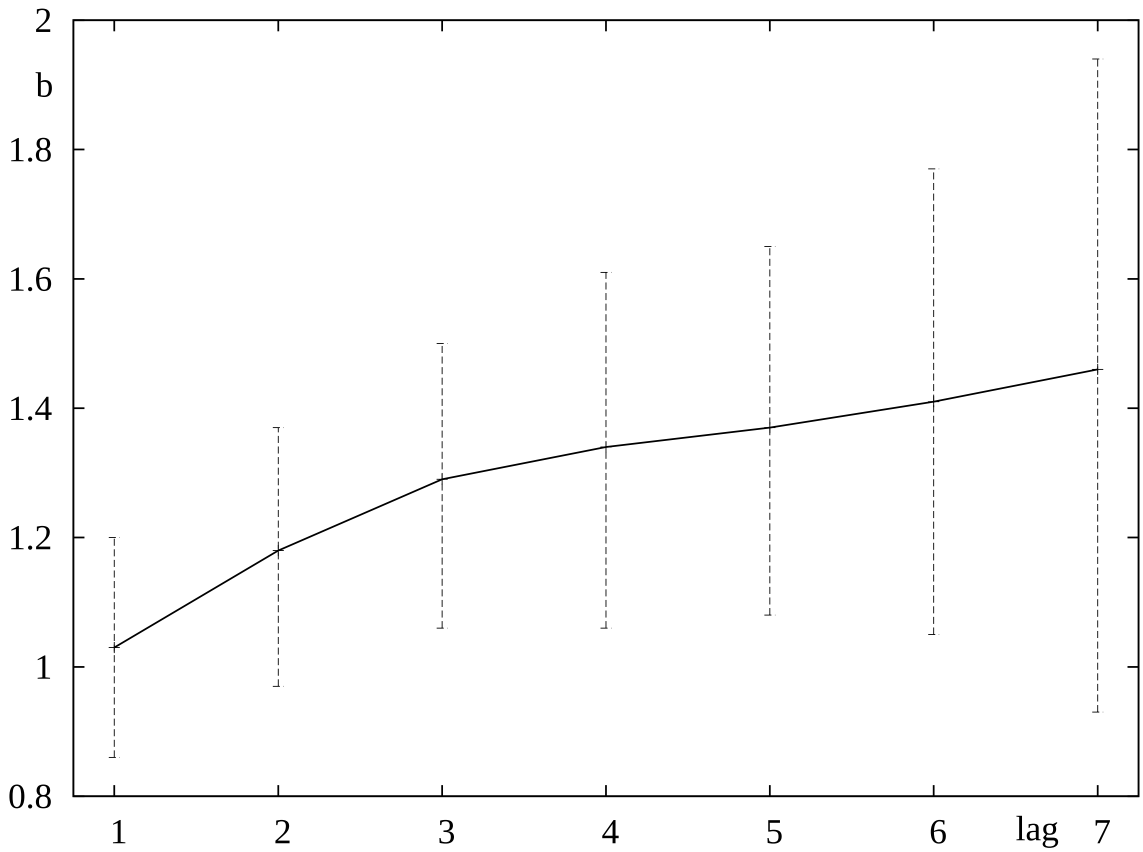

Figure 6: Estimated Subbotin shape parameter b of the growth rates distribution for different time horizons. The value reported is the average over all the sectors.

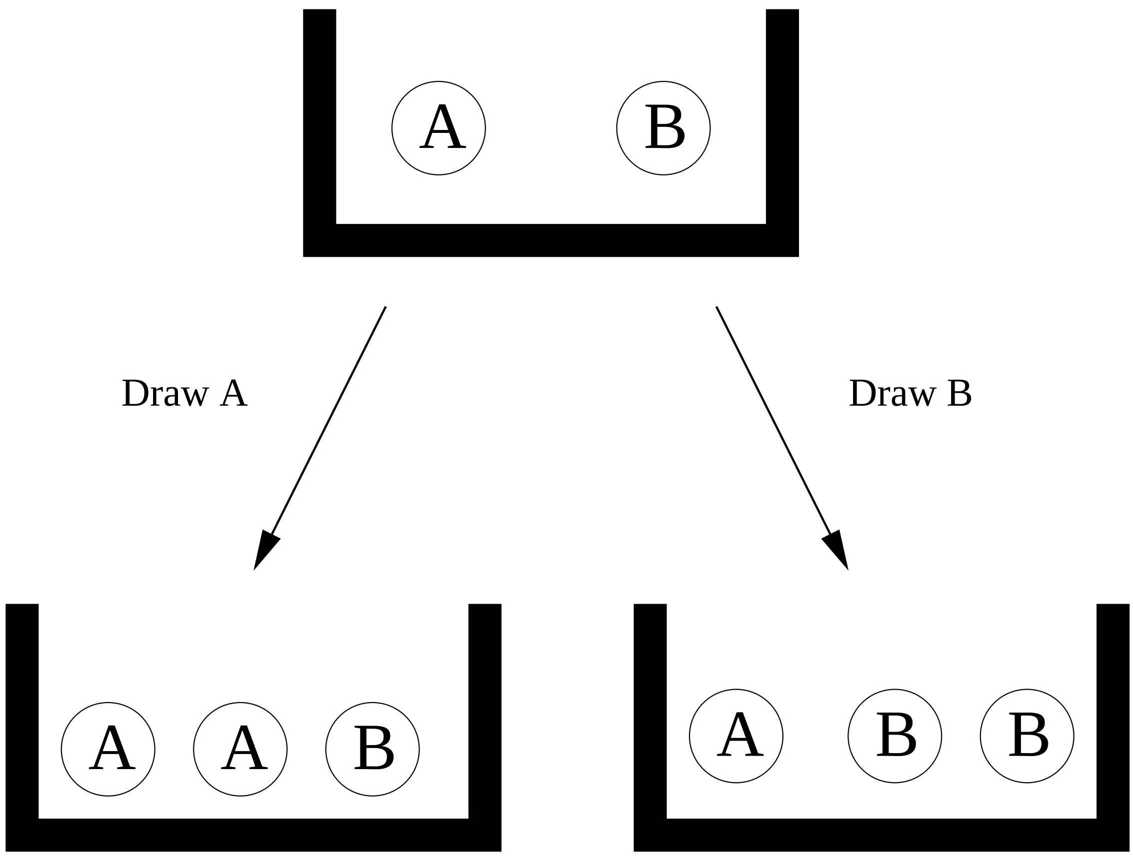

Figure 7: First drawn from a urn with two colors labeled A and B. After the extraction the state of the urn depends on which color was drawn.

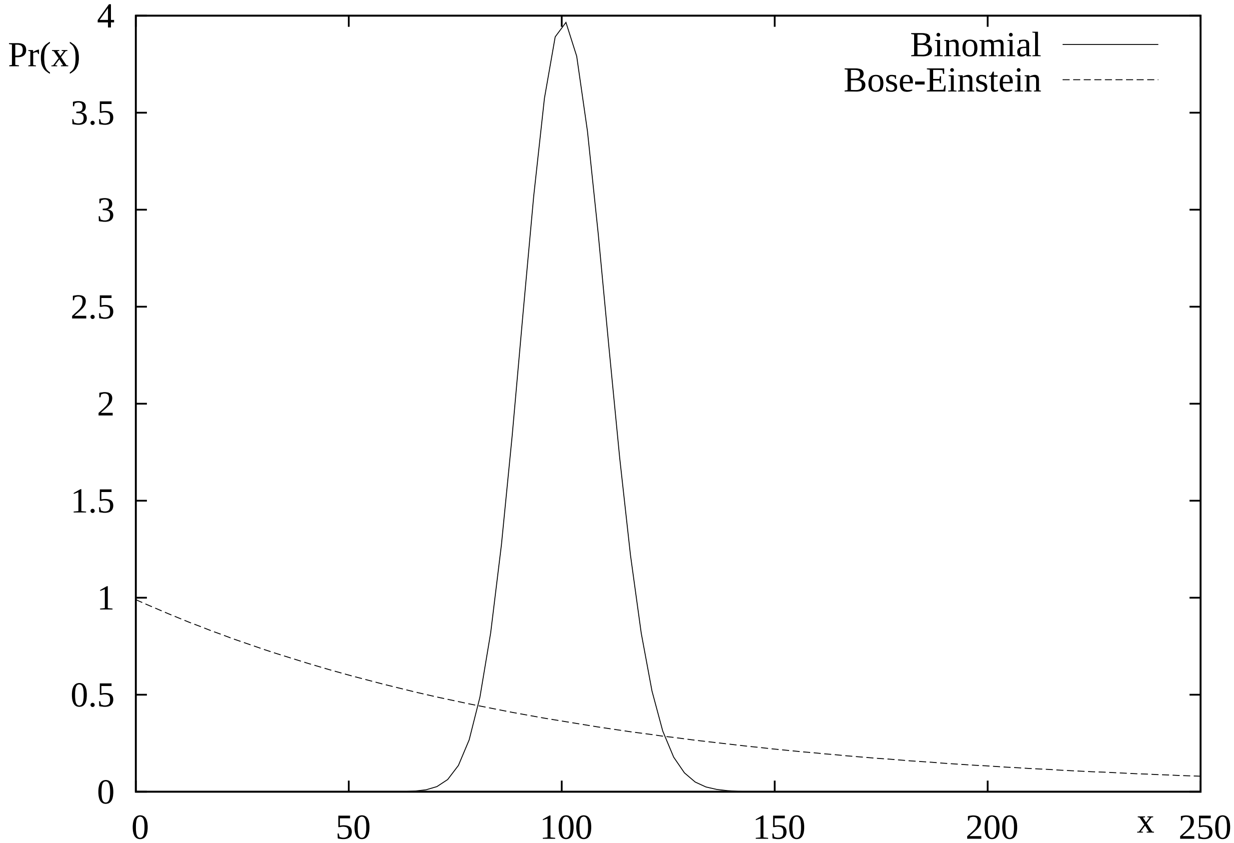

Figure 8: The Bose-Einstein probability distribution (9) of the number of opportunities per firm together with the corresponding Binomial distribution with $N = 100$ and $M = 10,000$.

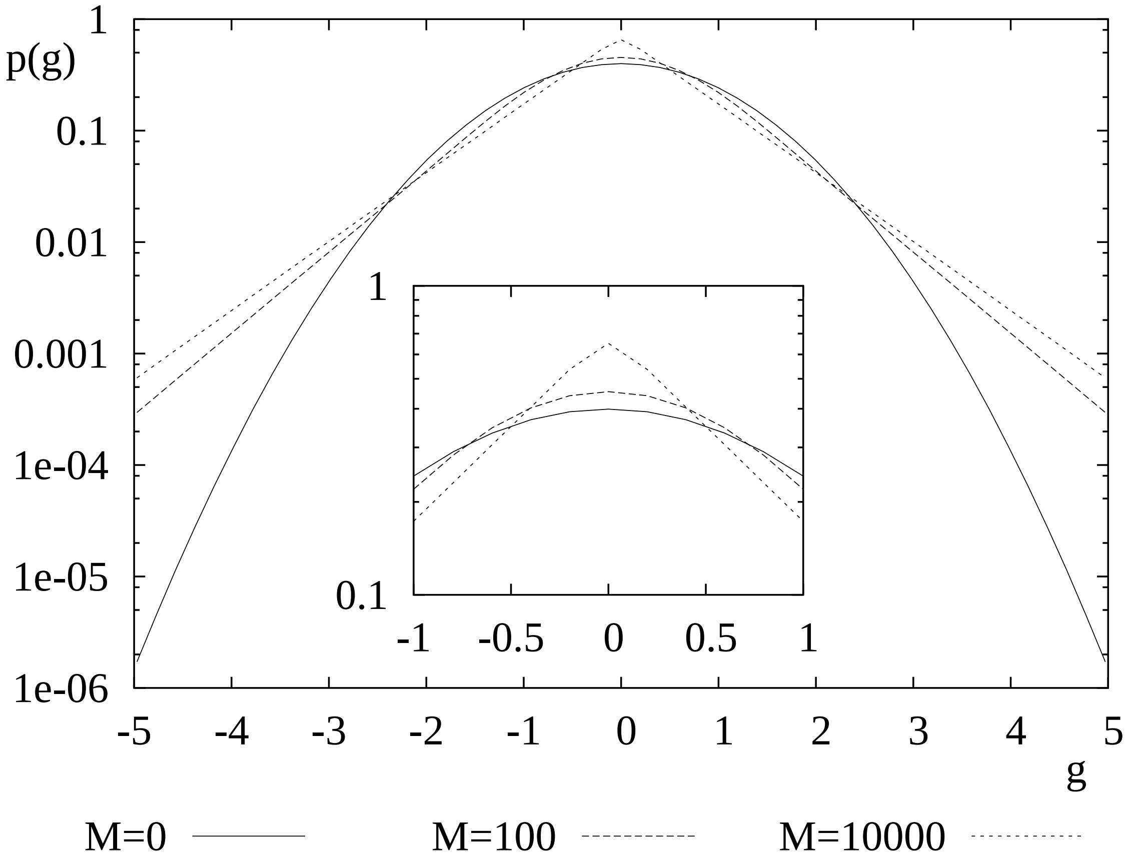

Figure 9: Growth rates probability density for $N = 100$ and different values of M. The inset shows a particular of the central region.

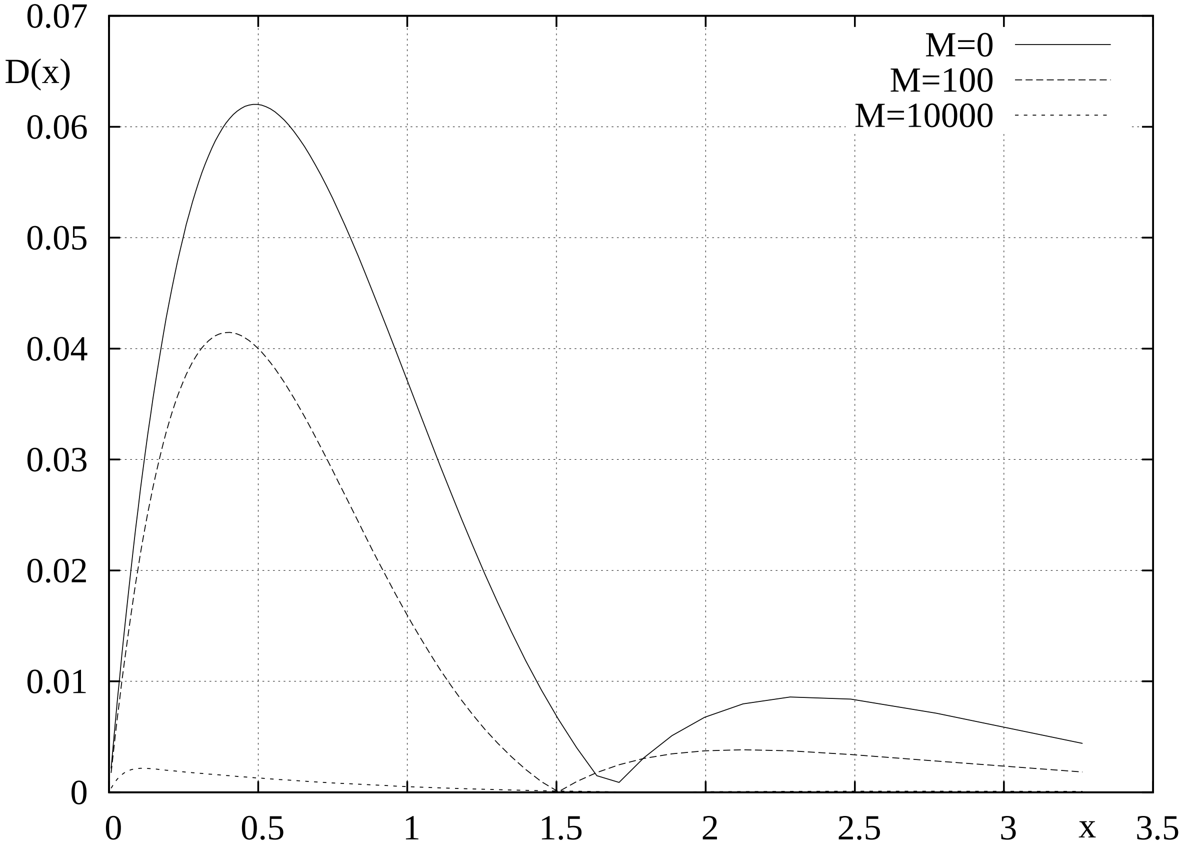

Figure 10: Absolute deviation $|F_{\rm model}(g) - F_{\rm L}(g)|$ as a function of g for $N = 100$ and different values of M.

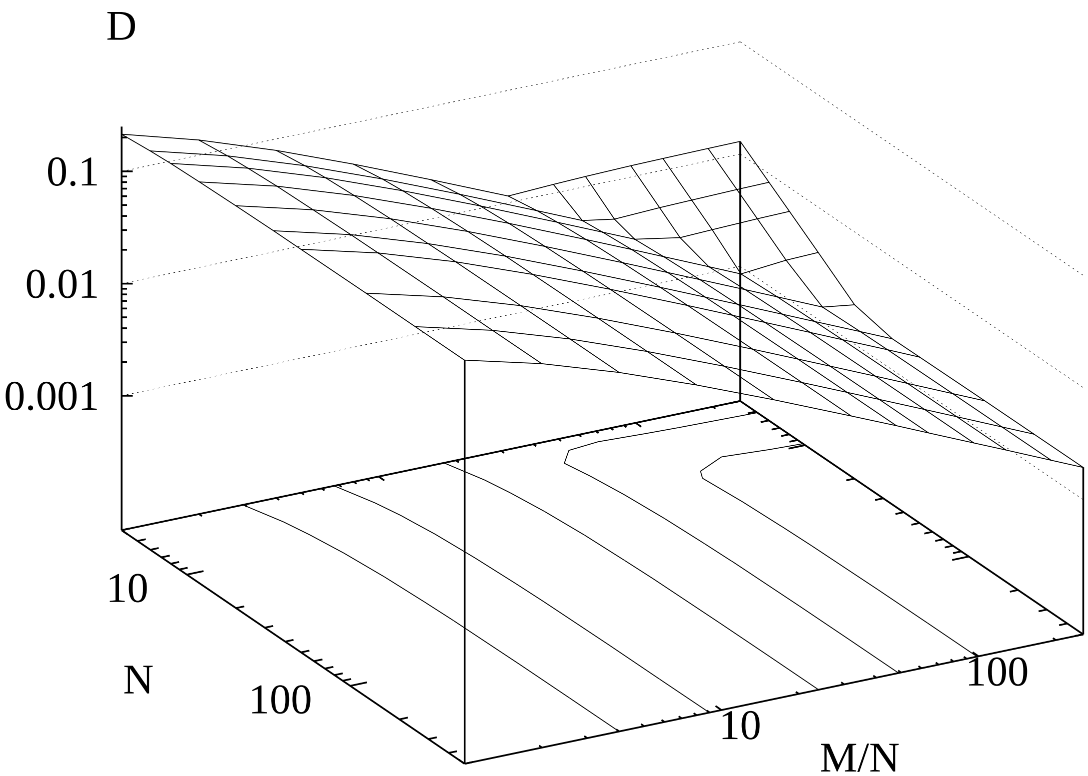

Figure 11: The maximal deviation D of the model density from a Laplace is shown as a function of the number of firms N and the average number of micro-shocks per firm M/N. Micro-shocks are normally distributed.

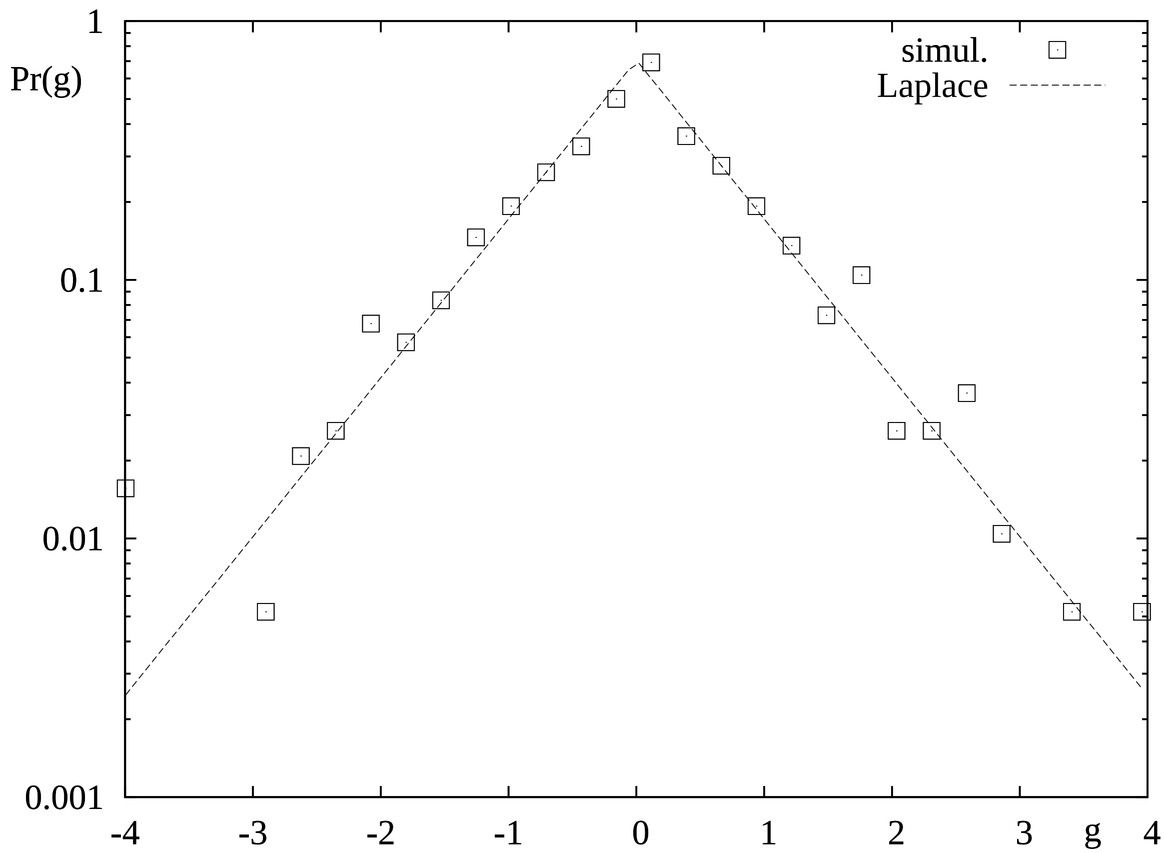

Figure 12: Growth rates probability density simulated with $\lambda = M/N = 16$, $N = 100$ and normally distributed micro-shocks. A pool of 7 independent realizations in considered. The theoretical Laplace density with unit variance ($a = 1/\sqrt{2}$) is also shown.

Ateco code	Sector	# of Firms	Autocorr. Coef.	Autocorr. Std Err.	Parameter b Coef.	Parameter b Std Err.	Parameter a Coef.	Parameter a Std Err.
151	Production, processing and preserving of meat	114	-0.15	0.08	0.83	0.05	0.089	0.004
155	Dairy products	85	-0.17	0.09	0.91	0.07	0.080	0.004
158	Production of other foodstuffs (brad, sugar, etc...)	157	-0.11	0.07	0.89	0.05	0.097	0.004
159	Production of beverages (alcoholic and not)	94	0.21	0.08	0.88	0.06	0.108	0.006
171	Preparation and spinning of textiles	154	0.02	0.07	1.19	0.07	0.142	0.005
172	Textiles weaving	171	-0.01	0.06	1.12	0.06	0.122	0.004
173	Finishing of textiles	181	0.13	0.06	1.11	0.06	0.107	0.004
175	Carpets, rugs and other textiles	90	-0.13	0.09	1.02	0.08	0.118	0.006
177	Knitted and crocheted articles	162	-0.09	0.07	0.97	0.05	0.124	0.005
182	Wearing apparel	379	0.01	0.05	0.92	0.03	0.120	0.003
191	Tanning and dressing of leather	87	0.04	0.09	1.12	0.09	0.140	0.007
193	Footwear	245	-0.06	0.05	1.12	0.05	0.150	0.004
202	Production of plywood and panels	52	-0.09	0.11	0.98	0.09	0.104	0.007
203	Wood products for construction	59	-0.28	0.11	0.94	0.08	0.105	0.007
205	Production of other wood products (cork, straw, etc...)	56	0.18	0.11	1.31	0.13	0.106	0.006
211	Pulp, paper and paperboard	46	-0.37	0.12	0.89	0.09	0.120	0.009
212	Articles of paper and paperboard	180	-0.19	0.06	0.93	0.05	0.103	0.004
221	Publishing	72	-0.11	0.10	0.62	0.05	0.079	0.005
222	Printing	199	-0.03	0.06	1.25	0.07	0.108	0.003
241	Production of basic chemicals	80	-0.17	0.09	0.88	0.07	0.114	0.006
243	Paints, varnishes, printing inks and mastics	58	-0.07	0.11	1.05	0.10	0.080	0.005
244	Pharmaceuticals, medicinal chemicals and botanical products	97	0.07	0.08	0.91	0.06	0.117	0.006
245	Soap and detergents, cleaning and toilet preparations	46	0.35	0.12	0.99	0.10	0.098	0.007
246	Other chemical products	51	-0.04	0.11	0.80	0.07	0.108	0.008
251	Rubber products	87	0.04	0.09	0.93	0.07	0.097	0.005
252	Plastic products	352	-0.12	0.04	0.95	0.04	0.113	0.003
261	Glass and glass products	87	-0.11	0.09	1.08	0.08	0.099	0.005
262	Ceramic goods not for construction	59	0.26	0.11	1.20	0.11	0.107	0.006
263	Ceramic goods for construction	91	0.09	0.08	1.04	0.08	0.109	0.005
264	Bricks, tiles and construction products in baked clay	84	-0.03	0.09	1.17	0.09	0.103	0.005
266	Articles in concrete, plaster and cement	141	-0.21	0.07	0.86	0.05	0.160	0.007
267	Cutting, shaping and finishing of stone	69	-0.03	0.10	1.19	0.10	0.116	0.006
273	First processing of iron and steel	82	0.05	0.09	0.84	0.06	0.126	0.007
275	Casting of metals	125	-0.13	0.07	0.99	0.06	0.116	0.005
281	Structural metal products	156	-0.18	0.07	1.26	0.07	0.185	0.007
284	Forging, pressing, stamping and roll forming of metal	132	-0.07	0.07	1.13	0.07	0.126	0.005
285	Treatment and coating of metals	182	-0.13	0.06	1.01	0.05	0.135	0.005
286	Cutlery, tools and general hardware	149	-0.07	0.07	1.01	0.06	0.125	0.005
287	Other fabricated metal products	265	-0.18	0.05	0.86	0.04	0.107	0.003
291	Machinery for the production and the use of mechanical power	224	-0.02	0.05	0.93	0.04	0.121	0.004
292	Other general purpose machinery	199	-0.22	0.06	1.04	0.05	0.158	0.005
293	Agricultural and forestry machinery	54	-0.34	0.11	0.96	0.09	0.139	0.009
294	Machine tools	114	-0.11	0.08	1.06	0.07	0.170	0.007
295	Other special purpose machinery	424	-0.24	0.04	1.06	0.04	0.185	0.004
297	Domestic appliances not elsewhere classified	59	-0.07	0.11	1.19	0.11	0.108	0.006
311	Electric motors, generators and transformers	71	-0.01	0.10	1.01	0.08	0.131	0.007
312	Manufacture of electricity distribution and control equipment	70	-0.16	0.10	0.84	0.07	0.149	0.009
316	Electrical equipment not elsewhere classified	91	-0.11	0.09	0.97	0.07	0.137	0.007
322	TV and radio transmitters and lines for telephony and telegraphy	44	-0.15	0.12	0.92	0.10	0.175	0.013
332	Measure, control and navigation instruments	51	0.09	0.11	1.20	0.12	0.119	0.008
342	Production of bodies for cars, trailers and semitrailers	50	-0.07	0.11	1.09	0.11	0.153	0.010
343	Production of spare parts and accessories for cars	125	-0.09	0.07	1.09	0.07	0.137	0.006
361	Furniture	444	-0.02	0.04	1.04	0.03	0.121	0.003
362	Jewelry and related articles	84	0.05	0.09	1.41	0.12	0.163	0.008
366	Miscellaneous manufacturing not elsewhere classified	68	0.10	0.10	1.14	0.10	0.117	0.007

Table 1: Summary table of the 55 sectors under analysis. For each sector are reported the estimated a and b parameters together with the growth rates autocorrelation coefficient.